

unine

UNIVERSITÉ DE
NEUCHÂTEL

Institut des sciences
du langage et de la
communication

Maud Dubois, Alain Kamber
& Carine Skupien Dekens (Eds)

L'enseignement de l'orthographe
en FLE

L'enseignement de l'orthographe en FLE

T R A

TRAVAUX NEUCHÂTELOIS DE LINGUISTIQUE

N E I

TRA
NEI

TRANEL (Travaux neuchâtelois de linguistique)

La revue TRANEL fonctionne sur le principe de la révision par les pairs. Les propositions de numéros thématiques qui sont soumises au coordinateur sont d'abord évaluées de manière globale par le comité scientifique. Si un projet est accepté, chaque contribution est transmise pour relecture à deux spécialistes indépendants, qui peuvent demander des amendements. La revue se réserve le droit de refuser la publication d'un article qui, même après révision, serait jugé de qualité scientifique insuffisante par les experts.

Responsable de la revue

Gilles Corminboeuf

email: gilles.corminboeuf@unine.ch

Comité scientifique de la revue

Marie-José Béguelin, Simona Pekarek Doehler, Louis de Saussure, Geneviève de Weck (Université de Neuchâtel)

Secrétariat de rédaction

Florence Rohrbach, Revue Tranel, Institut des sciences du langage et de la communication, Université de Neuchâtel, Espace Louis-Agassiz 1, CH-2000 Neuchâtel

Les anciens numéros sont également en accès libre (archive ouverte / open access) dans la bibliothèque numérique suisse romande Rero doc. Voir rubrique "Revue": <http://doc.rero.ch/collection/JOURNAL?ln=fr>

Abonnements

Toute demande d'abonnement ou de numéro séparé est à adresser à:

Revue Tranel, Université de Neuchâtel, Espace Louis-Agassiz 1, CH-2000 Neuchâtel

Tél.: ++41(0)32 718 16 90

Fax: ++41(0)32 718 17 01

email: revue.tranel@unine.ch

Tarifs

Abonnement annuel (2 numéros)	Suisse: CHF 51.00	Etranger: € 34.80
Numéro séparé	Suisse: CHF 27.00	Etranger: € 18.40
Numéro double	Suisse: CHF 40.00	Etranger: € 27.30

Paielement

Suisse: CCP 20-4130-2 – Université, Fonds de tiers, 2000 Neuchâtel (réf: U.00695)

Etranger: Compte: 290-500080.05H auprès d'UBS SA, 2000 Neuchâtel (Switzerland)

[Code Swift: UBSWCHZH20A] [IBAN: CH53 0029 0290 5000 8005 H]

© Institut des sciences du langage et de la communication, Université de Neuchâtel, 2011

Tous droits réservés

ISSN 1010-1705

Table des matières

- Maud DUBOIS, Alain KAMBER & Carine SKUPIEN DEKENS
Avant-propos ----- 1-6
- Daniel LUZZATI
L'orthographe dans l'élaboration des niveaux
de référence pour le français ----- 7-19
- Martha MAKASSIKIS & Jean-Christophe PELLAT
Les étudiants natifs et allophones face à
l'orthographe française: le cas des
homonymes ----- 21-48
- Saliha AMOKRANE
Du lien entre la maîtrise du système
phonologique et les compétences
orthographiques ----- 49-61
- Danièle GEFFROY KONŠTACKÝ & Sylva NOVÁKOVÁ
De la phonologie à la graphie vers leur
synergie: une expérience tchèque ----- 63-72
- Claire FONDET & Fabrice JEJCIC
OrthoFonic: un projet de didacticiel pour
l'apprentissage de l'orthographe française ----- 73-92
- Christian SURCOUF
L'enseignement et l'apprentissage de la
conjugaison en FLE: comment réduire les
difficultés engendrées par l'orthographe ----- 93-112
- Jeanne GONAC'H & Clara MORTAMET
Pratiques orthographiques en français
d'étudiants étrangers: le cas d'étudiants
hispano-américains et afghans ----- 113-127

▪ Jean-Marc DEFAYS & Frédéric SAENEN De l'incidence du cotexte sur les choix orthographiques en FLE: étude de cas -----	129-146
▪ Claude GRUAZ Pour une méthode active de remédiation orthographique pour des apprenants allophones -----	147-157
▪ Marie-Claude LE BOT & Elisabeth RICHARD Pour une analyse morphologique des productions écrites d'élèves FLS -----	159-172
Adresses des auteurs -----	173-174
Comité de lecture -----	175

Avant-propos

Maud DUBOIS, Alain KAMBER & Carine SKUPIEN DEKENS

Institut de langue et civilisation françaises, Université de Neuchâtel

L'apprentissage de l'orthographe constitue à l'évidence une difficulté majeure pour les apprenants du français allophones. Par ailleurs, l'orthographe étant un critère de sélection sur le marché du travail et un facteur de stigmatisation sociale, on peut légitimement considérer l'acquisition de cette compétence comme prioritaire pour des étudiants allophones qui peuvent être amenés à poursuivre un cursus universitaire ou à exercer une activité professionnelle dans un pays francophone.

Il semble que la pratique de l'enseignement de l'orthographe en FLM et en FLE n'ait pas toujours su tenir compte des spécificités du public visé, considérant que les zones orthographiques problématiques étaient les mêmes pour les francophones et les allophones. Ainsi, dans un certain nombre de manuels, les règles étaient présentées par difficulté croissante, puis venaient des listes d'exception, et des exercices. Dans la mesure où l'on élabore un enseignement spécifique pour allophones, on se trouve confronté à des choix méthodologiques. Depuis les travaux de Catach (1980), on prend davantage en considération la structure complexe de la langue à enseigner, mettant l'accent sur l'un ou l'autre des systèmes (phonogrammique, morphogrammique et logogrammique) de l'orthographe française. Les travaux présentés ci-dessous viennent enrichir cette approche, en la nuancant et en la complétant, et en apportant des éléments nouveaux, notamment la mise en valeur de la langue maternelle et des connaissances préalables des apprenants.

Adopter une approche plurisystémique

Si l'on veut décrire la compétence orthographique d'apprenants allophones d'une part, et permettre à ceux-ci d'améliorer leurs compétences d'autre part, il faut impérativement intégrer des facteurs tels que le rapport – ou la tension – entre la phonie et la graphie, les connaissances de la morphosyntaxe et de la sémiographie, mais également les compétences extralinguistiques.

Daniel LUZZATI explique pourquoi la maîtrise de l'orthographe est considérée, dans les référentiels se fondant sur les descripteurs du CECR, comme une composante des compétences de l'écrit de l'apprenant, et non

comme une compétence spécifique autonome. A ce jour, les descriptifs des quatre niveaux A1, A2, B1 et B2 (Luzzati 2004, 2007, 2008 et 2011) sont prêts et présentent les objectifs possibles d'apprenants du français. La hiérarchisation des niveaux de compétence intègre à la fois le rapport graphèmes / phonèmes, l'homophonie lexicale, l'orthographe grammaticale et la morphologie verbale. Les niveaux C en revanche s'avèrent beaucoup plus difficiles à décrire, la maîtrise orthographique impliquant, à l'instar du français langue maternelle, des variations d'ordre individuel. Ce manque d'homogénéité ne permet pas non plus le recours aux descripteurs existant pour le français langue maternelle. L'auteur insiste sur le fait que, loin de la sacralisation dont l'orthographe fait l'objet en français, un niveau C se distingue avant tout par des compétences réflexives, la maîtrise d'outils et la capacité de l'apprenant à mobiliser les compétences qu'il a acquises.

C'est à partir de la complexité du système orthographique français, "pluri-système" ou "système de systèmes" (Catach, 1980) que Martha MAKASSIKIS & Jean-Christophe PELLAT exposent leur recherche. L'articulation entre les principes sémiographique (particulièrement importants en français avec les lettres muettes, par exemple) et phonographique, propre à chaque langue, devrait être prise en compte dans les manuels d'orthographe, ce qui n'est pas toujours le cas pour ceux du français. Selon les auteurs, le principe sémiographique est souvent relégué au second plan. De ce point de vue, deux méthodes d'enseignement de l'orthographe FLE sont examinées de façon critique, *L'orthographe progressive du français, niveau débutant* (Chollet & Robert, 2004), et le *Manuel d'orthographe pour le français contemporain* (Skupien Dekens, Kamber & Dubois, 2011). Les auteurs présentent ensuite une enquête visant à comparer les performances orthographiques des étudiants FLE à celle des étudiants FLM, en particulier en ce qui concerne les homonymes. Il en ressort essentiellement que les étudiants francophones commettent plus d'erreurs logogrammiques et étymologiques que les étudiants allophones qui, eux, commettent plus d'erreurs phonogrammiques. Les auteurs concluent à la nécessité pour les concepteurs de manuels d'orthographe (FLM ou FLE) de présenter les dimensions phonographique et sémiographique de manière "à la fois séparée et combinée".

Tenir compte de la phonologie des langues maternelles

Les apprenants ayant une langue maternelle dont le système phonologique est éloigné du français devraient pouvoir bénéficier avant toute chose d'un enseignement de phonétique qui leur permettrait dans un deuxième temps d'acquérir l'orthographe du français. Les expériences menées par les auteurs des deux articles suivants parviennent à la même conclusion: sans une bonne maîtrise de la phonologie du français, l'orthographe ne peut être apprise efficacement.

A partir d'un corpus de productions écrites d'étudiants algériens en licence de français comprenant trois activités distinctes (dictée, rédaction et copie de texte), Saliha AMOKRANE examine si, à un niveau universitaire où elle n'est plus enseignée, l'orthographe est réellement maîtrisée et quels sont les problèmes rencontrés par les scripteurs. L'un des principaux résultats de cette recherche est qu'un très grand nombre d'erreurs sont imputables à une mauvaise maîtrise du système phonologique, et ce pour les trois types de tâches demandées. Ces erreurs concernent essentiellement les phonèmes vocaliques, qui diffèrent grandement entre le français et l'arabe. L'auteure en déduit qu'un enseignement approprié de phonétique devrait être dispensé aux étudiants arabophones au-delà de ce qui est prescrit dans les programmes scolaires (Amokrane, 2007) et qu'il devrait exercer aussi bien la discrimination auditive que la correction phonétique en production orale.

Après avoir décrit les systèmes phonologiques et prosodiques du français et du tchèque, Danièle GEFFROY KONŠTACKÝ & Sylva NOVÁKOVÁ montrent comment les particularités du français posent des problèmes aux locuteurs tchécoslovaques, particulièrement dans la chaîne parlée: difficulté à reconnaître des mots et à les segmenter, difficultés dans le passage d'un graphème à plusieurs phonèmes, (notamment avec le *e* caduc). Une expérience de terrain (sous la forme d'une dictée en première année de pédagogie à l'Université) montre en effet que les problèmes principaux sont dus à une mauvaise perception de la phonologie du français. En conséquence, les auteures proposent un travail en amont permettant, dès le début de l'apprentissage, l'acquisition systématique des compétences phonético-phonologiques. Elles proposent également de ne pas se contenter d'enseigner les mots isolément, mais de favoriser aussi la prosodie. Elles présentent brièvement leur manuel, *Entre nous* (Nováková et al., 2010 et 2011), dans lequel l'apprentissage commence de façon exclusivement audio-orale.

Utiliser la transcription phonétique

Indépendamment de toute référence à la langue maternelle de l'apprenant, l'enseignement de l'orthographe aurait tout à gagner d'un recours systématique à l'alphabet phonétique. Ceci afin de permettre un enseignement plus complet, qui intègre par exemple les phénomènes de liaisons ou les lettres muettes, ou pour donner accès à un système plus régulier et plus simple, celui de la morphologie verbale de l'oral.

Claire FONDET & Fabrice JEJCIC présentent leur logiciel *OrthoFonic* construit sur une structuration scientifique du vocabulaire, notamment en fonction de listes de fréquence, construites à partir d'enquêtes ou de corpus informatisés, et qui font apparaître une grande concentration du

vocabulaire d'usage très courant. Ainsi, il s'agit de privilégier l'enseignement des formes les plus fréquentes, qui constituent un "noyau" du français. L'orthographe rectifiée y est privilégiée. Les différentes listes sont présentées, avec leur transcription phonétique en Alfonic (voir Martinet et al., 1983 et Martinet, 1989) et en API, ainsi qu'avec les différentes utilisations didactiques possibles, grâce au codage spécifique du logiciel. Le codage Alfonic est explicité, ainsi que les fonctions des divers champs de la base de données. Les auteurs donnent ensuite des exemples concrets d'utilisation en classe de la base de données, dans le cadre d'un apprentissage spécifique notamment des graphies des finales, des liaisons, des homophones, des accords des adjectifs ou de la morphologie des noms. Ils appellent à une collaboration entre chercheurs et pédagogues, avec la participation de spécialistes en traitement automatique des langues.

Partant du constat que l'acquisition des formes verbales écrites du français est rendue plus difficile pour un apprenant allophone par le fait que ce dernier doit à la fois intégrer les formes et leur orthographe, Christian SURCOUF montre l'avantage que constitue l'apprentissage du système verbal par ses formes orales transcrites en API: une comparaison des codes écrit et oral met en évidence que la complexité orthographique du système de conjugaison des verbes – accentuée par la manière de le présenter dans les ouvrages de référence traditionnels – masque les régularités de l'oral (voir Pouradier Duteil, 1997). L'auteur propose une classification des verbes ainsi que des tableaux de formes verbales basés sur le code phonique susceptibles d'être utilisés comme une passerelle vers l'apprentissage de l'orthographe.

Bâtir sur les connaissances préalables

Une autre tendance qui se dégage des contributions de ce volume est l'importance de la langue maternelle et des connaissances préalables des apprenants.

Sur la base d'un corpus de textes d'apprenants de FLE produits dans deux contextes distincts – Amérique hispanophone et Afghanistan –, Jeanne GONAC'H et Clara MORTAMET analysent les "zones orthographiques problématiques" de ces apprenants, adoptant l'hypothèse selon laquelle la ou les langue(s) présente(s) dans leur répertoire peuvent influencer les erreurs commises (voir aussi Mortamet & Gonac'h, 2011), quantitativement différentes de celles de scripteurs natifs. Il s'agit ici de comparer les performances d'apprenants dont la langue maternelle est proche du français – l'espagnol – ou au contraire très éloignée – le dari – pour des erreurs concernant le marquage du genre et du nombre ainsi que des erreurs phonographiques. Les auteures montrent que si des points communs peuvent être mis au jour, il existe bel et bien des différences

spécifiques liées au système propre des langues sources concernées.

A partir d'un corpus de productions écrites de vingt étudiants de FLE de niveau B1 et de nationalités diverses, Jean-Marc DEFAYS et Frédéric SAENEN examinent les divers facteurs qui interviennent dans l'acquisition de l'orthographe. L'influence de la langue maternelle (voir Castelloti, 2005), le rôle que peuvent jouer certaines langues relais (à l'instar de l'anglais), mais également la capacité à prendre en compte le cotexte d'un mot sont des éléments cardinaux de la performance écrite. En opérant un test en deux phases distinctes (mots isolés / mots en contexte), les auteurs montrent ainsi que la contextualisation d'un mot sur un axe syntagmatique peut aider les apprenants à l'orthographier correctement, mais qu'elle reste insuffisante dans certains cas, notamment en ce qui concerne les marques du pluriel. Ces erreurs indiquent que le sens de la phrase reste opaque pour les apprenants, mais aussi que l'articulation simultanée de l'orthographe d'usage et de l'orthographe grammaticale constitue une difficulté majeure pour eux.

Claude GRUAZ propose une approche de l'enseignement de l'orthographe du français à des apprenants étrangers qui valorise le plus possible les compétences orthographiques et les automatismes déjà en place. L'analyse des écarts, sur la base de textes produits par les apprenants, permet d'identifier leurs sources (erreur de prononciation, problème purement graphique, aspect morphologique ou de marquage spécifique), d'établir un profil orthographique et de proposer un processus de remédiation à l'apprenant (voir Crefor Haute-Normandie, 2009). L'auteur insiste sur la nécessité de mener une réflexion avec l'apprenant sur sa représentation de l'orthographe et de ses propres difficultés, puis sur la mise en évidence de régularités dans le système orthographique, à partir de ses connaissances préalables. Le rôle du formateur est d'entraîner l'apprenant à manipuler la langue (par substitution et analogie), pour qu'il puisse réactiver certaines connaissances et découvrir les grandes régularités du système.

Examinant un corpus de textes produits par des élèves majoritairement turcophones d'un collège breton, Marie-Claude LE BOT et Elisabeth RICHARD ont pour objectif de décrire et d'expliquer les erreurs commises par des scripteurs que leur performance met en grande difficulté scolaire alors même que leur maîtrise du code oral est satisfaisante. L'analyse s'appuie sur l'hypothèse que les erreurs ne sont pas seulement des écarts fautifs par rapport à une norme, mais qu'elles révèlent des stratégies cohérentes dans le cadre d'un apprentissage en cours (Cogis, 2002). Les auteures montrent que cette approche permet de réduire l'image négative véhiculée par ces textes et de mettre en évidence les acquis de ces élèves dans le code graphique, mais également de mieux cerner leurs lacunes, notamment la maîtrise de la morphologie verbale et nominale (Besse & Porquier, 1984).

A travers des approches diverses, les travaux réunis dans ce volume démontrent la nécessité de changer la perspective de l'enseignement de l'orthographe FLE, et de mettre l'apprenant au centre du processus, avec sa langue, sa phonologie, ses stratégies et ses tentatives d'appropriation de l'orthographe française.

Bibliographie

- Amokrane, S. (2007): Pour améliorer la maîtrise de l'orthographe en FLE/FLS... renforcer l'enseignement de la phonétique. In: les Annales de l'université d'Alger, n°17 tome, 31-37.
- Besse H. & Porquier R. (1984): Grammaires et didactique des langues. Paris (Hatier).
- Castelloti, V. (2005): La Langue maternelle en classe de langue étrangère. Paris (CLE International).
- Catach, N. (1980): L'Orthographe française. Paris (Nathan).
- Chollet, I. & Robert, J.-M. (2004): Orthographe progressive du français, niveau débutant. Paris (CLE International).
- Cogis, D. (2002): Comment le genre graphique vient aux enfants. In: Hass, G. (dir): Apprendre, comprendre l'orthographe autrement de la maternelle au lycée, CRDP Bourgogne, 19-42.
- Crefor Haute-Normandie, coordination Mercier J.-P., direction Gruaz, C., collaboration Bottois, F., Chesnel, C. & Funkiewiez, F (2009): Recherche-action Orthographe et illettrisme.
(http://www.crefor-hn.fr/sites/default/files/Recherche_action_orthographe.pdf)
- Luzzati, D. (2004): Chapitre 8: "la matière graphique". In: Niveau B2 pour le français, un référentiel, 289-326, Beacco, J.-C., Bouquet, S. & Porquier R. (dir.). Paris (Didier).
- Luzzati, D. (2007): Chapitre 8: "la matière graphique". In: Niveau A1 pour le français, un référentiel, 139-155, Beacco, J.-C. & Porquier R. (dir.). Paris (Didier).
- Luzzati, D. (2008): Chapitre 8: "la matière graphique". In: Niveau A2 pour le français, un référentiel, 139-155, Beacco, J.-C. & Porquier R. (dir.). Paris (Didier).
- Luzzati, D. (2011): Chapitre 8: "la matière graphique". In: Niveau B1 pour le français: un référentiel, Beacco, J.-C., Blin, B., Houles, E., Lepage, S. & Riba, P. (dir.). Paris (Didier).
- Martinet, A. & J., Villard, J., Boyer, D. & Dominici, G. (1983): Vers l'écrit avec Alfonic. Paris (réédition 2006, Association RAPHAEL).
- Martinet, A. (1989): Alfonic au plus près de l'orthographe. In: Liaison Alfonic, fascicule 2, 7-12.
- Mortamet, C. & Gonac'h, J., (2011): Variation orthographique en français: le cas des non-natifs. In: Schnaffer, I. & Bertrand, O. (dirs.): Variétés, variations et formes du français. Paris (Les éditions de l'école Polytechnique).
- Nováková, S., Kolmanová, J., Geffroy Konšťáky, D. & Táborská, J. (2009): Le français ENTRE NOUS 1. Plzeň (Fraus).
- Nováková, S., Kolmanová, J., Geffroy Konšťáky, D. & Táborská, J. (2010): Le français ENTRE NOUS 2. Plzeň (Fraus).
- Pouradier Duteil, F. (1997): Le Verbe français en conjugaison orale. Francfort/Main (Peter Lang).
- Skupien Dekens, C., Kamber, A. & Dubois, M. (2011): Manuel d'orthographe pour le français contemporain. Neuchâtel (Alphil – Presses universitaires suisses).

L'orthographe dans l'élaboration des niveaux de référence pour le français

Daniel LUZZATI

LIUM, Université du Maine

In the framework of the CEFR, language versions for French devotes a chapter from A1 to "graphic material". These chapters are designed according to phonology, vocabulary –especially homophones –, grammatical words and verbal features. A1/A2 & B1/B2 are now completed. Thus, it becomes interesting to ask a few questions:

- Is the title "matière graphique" relevant?
- Prejudgements (plan, verbal morphology by basis...) are they timely?
- Is the hierarchy of skill levels working with spelling?
- Is the skill levels succession expected by the CEFR working with spelling?
- What could be C reference levels?

Introduction

Pour l'élaboration des chapitres consacrés à la *Matière graphique* du français dans les référentiels issus du CECR¹, nous sommes parti de l'idée que les difficultés de l'écriture du français peuvent s'expliquer par un principe relativement simple: les mots latins, qui ont conservé le même nombre de syllabes en italien, espagnol, portugais, roumain... se sont systématiquement réduits en français. En général, l'accent est demeuré au même endroit, mais les syllabes postérieures (et parfois antérieures) ont disparu ou se sont amuïes, notamment à l'oral: *asinum* a donné *âne*, *regem* est devenu *roi*, *nominem* s'est réduit en *nom*; *caballus*, au lieu de donner *cavallo* (*caballo*, *cavalo*...) est devenu *cheval*, ou plutôt /ʃfal/ avec la disparition du schwa, qui induit en outre des assimilations². On aboutit ainsi à un recours massif, d'une part à des morphèmes grammaticaux liés qui ne s'entendent pas (*nie*, *nies*, *nient*), et d'autre part à une surabondance de mots/formes monosyllabiques homophones hétérographes (*nid*, *nids*, *ni*), en concurrence parfois avec des séquences de mots grammaticaux (*n'y*). On peut illustrer cette caractéristique du français à l'aide de la Figure 1.

¹ Dans Luzzati (2004, 2007, 2008, 2011) nous avons rédigé les chapitres correspondants dans les niveaux de références pour le français parus chez Didier (*Niveaux A1/A2/B1/B2 pour le français, un référentiel*).

² Indépendamment du fait qu'une forme *cavale* demeure en usage.

	Français	Anglais	Italien
Consonnes	17/15 ³	24	14
Voyelles	15/10 ⁴	20	5
Nombre de combinaisons-mots C/V	255/150	480	70
% combinaisons-mots vides ⁵	3,33%	35%	62%
Nombre moyen de combinaisons-mots ou "taux de confusion"	4,97	0,47	0,65

Fig.1: taux de confusion des séquences C/V en français, anglais et italien, issus de Luzzati (2010: 37)

En d'autres termes, l'écriture de l'anglais est difficile parce que la relation phonèmes/graphèmes est particulièrement complexe, du fait d'un système phonologique très riche, qui produit un nombre de combinaisons-mots important, dont plus du tiers reste inexploité. La complexité de l'écriture française est la conséquence à la fois d'une relation phonèmes/graphèmes complexe et d'une multiplication des homophones monosyllabiques qui, outre qu'elle va croissant du fait de l'émergence des voyelles opposées, impose une hétérographie nécessairement ardue, ce que dénote un "taux de confusion" près de 10 fois supérieur à celui de l'anglais ou de l'italien, avec un nombre très faible de combinaisons-mots inexploitées. En italien, l'écriture ne pose guère de problème, que ce soit pour les italophones ou pour les apprenants étrangers, dans la mesure où ce qui s'écrit s'entend et ce qui s'entend s'écrit, avec une faible proportion de mots monosyllabiques, et une majorité de combinaisons-mots inexploitées.

Pour le français, la maîtrise de l'orthographe ne pose pas les mêmes problèmes en FLE qu'en langue maternelle, à tel point qu'on peut penser que c'est essentiellement la maîtrise d'une langue étrangère qui pourrait aider les apprenants francophones à pallier certaines de leurs difficultés. Dès lors qu'on peut les traduire aisément, les homophones grammaticaux, seuls (*a/à, et/est/ai/ait/aie/aies/aient, ce/se, on/ont, son/sont, ses/ces/c'est/s'est/sait...*) ou en séquence (*la/l'a, les/l'est/l'ai, qui la/qui l'a/qu'il a/qu'il la/qu'il l'a...*), deviennent massivement transparents, avec accès direct au sens, sans avoir à passer nécessairement par une case grammaire. De ce fait ils constituent *a priori* moins une difficulté majeure en FLE que les morphèmes grammaticaux muets, ou bien que les homophones lexicaux, qui posent des problèmes considérables, en relation avec la maîtrise de leur sens.

³ L'existence des consonnes /ɲ/ et /ŋ/ étant discutable, le français standard comporte entre 15 et 17 consonnes.

⁴ Les 5 voyelles opposées (/ɛ/ et /œ/, /e/ et /ɛ/, /ɔ/ et /o/, /ø/ et /œ/, /a/ et /ɑ/) étant de moins en moins discriminées, le système vocalique actuel du français standard comporte entre 10 et 15 voyelles.

⁵ Nous entendons par "combinaison-mot C/V" les mots composés d'une consonne et d'une voyelle (*jeu, jet, gît, jus, joue, gens, jeun, jonc* par exemple), et les "combinaisons-mots vides" sont les séquences C/V qui ne correspondent à aucun mot possible (/ʒo/ ou (/ʒɔ/ par exemple).

On comprend dès lors que, dans le cadre de la construction de référentiels pour le français, il ait fallu construire un chapitre spécifique, consacré à l'orthographe. Cela posait un certain nombre de problèmes qu'il fallait résoudre dans le cadre contraint de la constitution de référentiels à partir des descripteurs du CECR (cf. annexe 1), les référentiels en question ayant à considérer la maîtrise de l'orthographe comme une des composantes des compétences de l'écrit, et non comme une compétence spécifique et autonome.

1. La *matière graphique*

Les chapitres en question s'intitulent *matière graphique*, ce qui pointe sur deux compétences distinctes: d'une part il s'agit de lire, à voix basse comme à voix haute (compétence orthoépique, commune à toutes les langues); d'autre part, il s'agit de maîtriser les pièges de l'écriture (compétence orthographique, dont la difficulté et le statut sont particuliers avec le français). Cela évoque ensuite l'ensemble des moyens utilisés pour écrire une langue. *Matière*, comme *matériau*, renvoie à quelque chose de massif, de plus naturel (de moins artificiel si on veut) que l'orthographe. La *matière graphique* est un ensemble dont la complexité est difficile à sonder, un ensemble à l'intérieur duquel on puise, sans nécessairement savoir pourquoi. Si on évoque une *matière graphique*, on lui présuppose en outre un "contraire", un "pendant" ou un "reflet": la *matière sonore*. On peut en somme penser l'écrit de deux manières différentes: comme la transposition de l'oral ou comme un système autonome, comme le font Anis et al. (1988) ou Catach (1991).

L'expression peut convenir à toutes les langues, à la différence d'*orthographe*. Ailleurs, on parle en général non pas de respect d'une norme, mais simplement de dire les lettres, d'épeler (*spelling*). Parfois, comme en espagnol ou en italien, le mot existe, s'écrit sans *h* (*ortografia*), et son usage correspond moins à une sacralisation de la norme graphique, qui ne fait pas de problème, qu'à l'idée de graphie et d'écriture. Il s'agit en somme essentiellement d'orthoépie et, à l'occasion seulement, d'orthographe. En français, il s'agit surtout d'orthographe, et le mot se décline: *orthographier*, *orthographie*, *orthographique* (famille qui ne demande qu'à s'agrandir, avec *orthographiquement*, *dysorthographie*...). Pour le français, l'*orthographe* est en somme bien davantage que l'*écriture* ou la *graphie*: matière, discipline, domaine, voire religion ou position idéologique, avec exercice fétiche à la clé, la dictée. Il était préférable, dans les référentiels, de ne pas renvoyer d'emblée à cette sacralisation, même si le terme est utilisé à l'intérieur des chapitres concernés.

Nous aurions pu certes recourir à l'expression *code écrit*, qui possède son pendant, le *code oral*. Elle présuppose qu'il existe une convention

exprimable et exécutable, et qu'en matière d'écriture, les choses peuvent être explicites, sinon régulières. Mais le *code* en question aurait eu l'inconvénient de porter une comparaison avec d'autres *codes*: le *code* de la route ou le *code* civil par exemple, ensemble de règles exprimables, constitutives de liens sociaux fondamentaux, dont le non respect, outre un danger éventuel, donne lieu à des infractions dûment réprimées. Leur maîtrise est réputée possible, quitte à recourir à une jurisprudence ample et complexe. Un *code écrit* du français serait ainsi présumé explicite et réductible à un ensemble de "règles", érigeant ce concept comme un sésame, alors qu'il peut apparaître comme fallacieux, notamment lorsqu'il repose sur une maîtrise illusoire de la grammaire. Qui plus est, le non respect d'un code écrit serait passible de sanctions, certes symboliques, mais protectrices d'un lien social ainsi mis à mal: une *faute* d'orthographe est certes en deçà d'un délit, mais néanmoins condamnable devant certains "tribunaux" scolaires. L'expression présuppose en somme l'existence d'une sentence orthographique, ce qui n'était pas le but recherché.

Nous aurions pu enfin utiliser *sémiographie* (étymologiquement, l'"écriture des signes"), qu'il faut comprendre comme la "représentation graphique du sens linguistique", tel que défini par Fayol & Jaffré (2008: 94), dont le pendant serait la *phonologie*. Mais on aurait alors placé les référentiels dans un champ théorique, alors qu'il s'agit d'outils destinés à un public plus large que les seuls linguistes.

2. Les partis-pris

Le premier parti-pris est celui des référentiels. Il s'agit non pas de construire un manuel d'orthographe mais de lister les objectifs possibles d'un apprenant-utilisateur pour les 4 niveaux actuellement traités (A1, A2, B1, B2), de façon à permettre à ceux qui ont à construire des *curricula*, des manuels, des cours... de disposer d'un outil adaptable en fonction d'objectifs et de publics divers. Le deuxième parti-pris a été celui de la méthode, qui a consisté à commencer par le niveau B2 pour s'attaquer ensuite dans l'ordre aux niveaux A1, A2 et B1, tout en traitant de la compétence graphique en relation avec les autres. La présentation peut en effet prendre un aspect réductionniste, pas la démarche ou les contenus. Les compétences lexicales doivent renvoyer aux "fonctions générales et spécifiques"; les compétences orthoépiques doivent s'appuyer sur le chapitre consacré à la *Matière sonore*; l'orthographe des formes verbales renvoie aux compétences grammaticales (ainsi, jusqu'à B2, le passé simple ne fait pas partie des compétences visées). Le troisième parti-pris consiste à partir des formes orales pour aller vers les formes graphiques. On présuppose ainsi que les compétences graphiques ne sont pas des compétences

autonomes, qu'elles sont secondes par rapport aux compétences orales (en réception comme en production), qu'elles doivent s'appuyer sur elles, et qu'elles ne doivent pas passer nécessairement par une maîtrise des concepts grammaticaux (français parlé et français grammatical normé ne sont pas nécessairement convergents). Le quatrième parti-pris consiste à accorder la primauté au sens sur la forme, aux signifiés sur les signifiants. A partir des chapitres organisateurs initiaux sur les "spécifications" et les "fonctions", qui font l'inventaire des compétences à maîtriser, on identifie quelles sont les formes à acquérir.

Nous avons systématiquement organisé les chapitres *Matière graphique* en 4 sections, en recourant cette fois au concept d'*orthographe*, au sens de norme graphique de référence, sans tenir compte des rectifications de 1990, dans la mesure où le nombre de points concernés, accents circonflexes sur *i* exceptés, est trop faible pour les niveaux du CECR qui nous occupent (de A1 à B2) :

- L'orthographe et la prononciation
- L'orthographe lexicale
- L'orthographe grammaticale
- L'orthographe des formes verbales

L'orthographe et la prononciation, c'est-à-dire la partie orthoépique, a été organisée par phonèmes vocaliques et consonantiques. Les glides (/w/, /ɥ/ et /j/) ont été intégrés à la suite des phonèmes vocaliques, qui traitent ensemble les voyelles opposées selon l'aperture, sauf /*ẽ*/ et /*œ*/: /*e*/ et /*ɛ*/, /*ɔ*/ et /*o*/, /*ø*/ et /*œ*/. Nous avons considéré 3 "niveaux" possibles, en distinguant les "graphies de base", les "graphies occasionnelles" et les "graphies rares" (en réduisant la classification à "graphies de base" et "autres graphies" en A1 et A2), ce qui se rapproche des descriptions de Gak (1976) ou Catach (1991) qui, d'une certaine manière, intégraient déjà l'idée de niveaux de compétence.

La compétence lexicale consiste en un inventaire des homophones hétérographes concernés par le niveau visé. Il ne s'agit pas en effet de reprendre les items recensés dans les chapitres consacrés aux "fonctions générales et spécifiques", mais de collationner ceux qui présentent des difficultés orthographiques. Nous avons eu recours à des moyens diacritiques, qui peuvent s'illustrer à partir des entrées suivantes: *haut*, {*os*}, *eau au*, *oh*, *ho*, (*aulx*) ou *pan*, *pend-s* <*pendre*, (*paon*), (*Pan*⁶). Une série

⁶ Ces notations réclament bien évidemment un mode d'emploi, dont voici quelques illustrations: "Les formes verbales sont indiquées avec des variations de désinence: *entre-es-ent* <*entrer* renvoie aux formes *j'entre*, *tu entres* et *ils entrent*; *entré+entraî-s-t-ent* <*entrer* renvoie aux formes *entré*, *entrés*, *entrée*, *entrées*, *j'entraî*, *j'entraîs*, *tu entraîs*, *il entraî* et *ils entraient* - Pour les adjectifs, les noms ou les participes, seules les formes du masculin singulier seront indiquées, même si le féminin et le pluriel sont souvent

ne sera présente que si 1 item au moins relève du niveau considéré et, s'il y en a plusieurs, elle apparaîtra plusieurs fois (aux entrées *au*, *eau* et *haut* par exemple). Elle peut toutefois inclure, entre parenthèses, des items relevant d'un niveau supérieur: l'importance d'une graphie se perçoit par rapport à l'ensemble du système. La majuscule signale les noms propres. Pour les formes verbales, il faut indiquer leurs variations ainsi que l'infinitif, ce qui se présente par exemple de la manière suivante: *dit*, {*dix*}, *di+dis+dit* <*dire*, ou *été*, *été+étais-t-ent* <*être*, (*étaie-es-ent* <*étayer*), (*étai*). Parfois enfin il existe, sinon plusieurs formes de certains items, du moins plusieurs prononciations possibles ({*os*} par exemple). Les listes concernent les mots que l'apprenant-utilisateur est censé reconnaître, entre autres parce qu'il est censé savoir s'il maîtrise ou non leur sens. La capacité à les produire (*vs* orthographier) constitue une cible variable, en fonction des utilisations des référentiels.

L'orthographe grammaticale est abordée autant que possible à partir de l'oral, en listant les morphèmes grammaticaux monosyllabiques et les séquences de mots grammaticaux ambiguës. La variabilité est abordée tout d'abord en listant les morphèmes invariables, puis les variations en genre et en nombre. Ces dernières sont abordées comme si la forme neutre était le masculin singulier, angle sinon le plus usuel, du moins le plus attendu. On évoque donc les marques, audibles et muettes, d'une part du féminin grammatical, d'autre part du pluriel grammatical.

Pour décrire la compétence morphologique relative aux formes verbales, et notamment aux radicaux, nous n'avons pas utilisé la subdivision traditionnelle en "groupes", évoquant simplement les *verbes* en *-er*, en *-ir* ou en *-r(e)*. Nous avons préféré la description par bases, issue de Pinchon & Couté (1980), qui permet de classer les verbes en fonction du nombre et de la distribution (suivant les temps et les personnes) des bases orales qu'ils comportent. Elle correspond à une approche de l'orthographe fondée sur les formes orales. Pour les utilisateurs des référentiels, elle permet également de s'appuyer sur certaines personnes davantage que sur d'autres, notamment en production: la personne 3 (*il/elle*) est évidemment d'un emploi beaucoup plus fréquent que la personne 4 (*nous*), à laquelle elle peut d'ailleurs se substituer, par l'intermédiaire de *on*. Elle fait en outre apparaître d'emblée que les verbes les plus faciles (*courir* par exemple) ne sont pas forcément les plus attendus, et que les verbes les plus difficiles (ceux qui comportent le plus de bases, c'est-à-dire les plus "irréguliers") sont aussi les verbes les plus courants (*être*, *avoir*, *savoir*, *vouloir*, *pouvoir*,

concernées: on indiquera *mur*, *mûr*, *mûre* (sans préciser *murs*, *mûrs* ou *mûres*), ou bien *entrée*, *entré+entrai-s-t-ent* <*entrer* (sans préciser *entrées* – pluriel du nom et féminin pluriel du participe – ou *entrés*) – On utilisera le codage par { } pour indiquer les mots à prononciation multiple (*six* ou *os* par exemple), et / pour indiquer les mots à graphies multiples (*clé/clefou* *soûl/saoul* ou *balaye/balaie* par exemple)."

aller, faire), en l'occurrence ceux qui peuvent fonctionner comme "auxiliaires".

3. La hiérarchisation des niveaux de compétence

La hiérarchisation des niveaux de compétences, c'est-à-dire la répartition section par section entre A1 et B2 est autant un résultat qu'un calcul. Elle comporte d'ailleurs quelques fluctuations ou incohérences. En ce qui concerne le rapport graphèmes/phonèmes, cela recoupe peu ou prou la logique des 77 "graphèmes de bases" de N. Catach (1991) dès A1, puis des quelque 133 graphèmes plus rares, en partie acquis entre A2 et B1, en considérant que cette acquisition concerne pleinement la réception et partiellement la production. La logique est double: il s'agit d'une part de respecter le fonctionnement du système (en intégrant simultanément par exemple les *m* devant *b/p* pour la graphie des nasales), et d'autre part la prééminence des signifiés, portée par les "fonctions générales et spécifiques". Ainsi, dès lors que *chien* ou *rien* apparaissent en A1, la graphie *en* pour [ɛ̃] s'impose, et la fréquence des mots visés importe davantage que celle des seuls graphèmes. En la matière, la hiérarchisation des niveaux de compétence suit une courbe non linéaire: 76 graphèmes en A1, 94 en A2, et 119 en B1 et 124 en B2. Autrement dit 61% des graphèmes-cibles recensés dans B2 doivent être reconnus dès A1, 75% (+14%) à compter de A2, 96% (+26%) à partir de B1, B2 n'en ajoutant que 4%. A titre d'exemple, on trouvera en Figure 2 un cas vocalique, et en Figure 3 un cas consonantique.

[ɛ̃]		exemples	niveaux
Graphie de base	in	<i>vin, inscrire</i>	A1
Graphies fréquentes	ain	<i>pain, vaincre</i>	
	ein	<i>plein, peindre, rein</i>	
	en	<i>chien, rien</i>	
	im	<i>impair</i>	
Graphies occasionnelles	aim	<i>faim</i>	A2
	yn	<i>synthèse</i>	B1
	ym	<i>symbole</i>	

Fig.2: exemple de hiérarchisation de la relation phonème-graphème pour un phonème vocalique

[k]		exemples	niveaux
Graphies de base	c	<i>curieux, couper</i>	A1
	qu	<i>chaque, quand</i>	
Graphies occasionnelles	k	<i>kilo</i>	
	q	<i>coq, cinq</i>	
	cc	<i>occasion, occuper</i>	A2
	ch	<i>orchestre, chœur</i>	
	ck	<i>ticket, bifteck</i>	B1
	cqu	<i>acquérir, grecque</i>	B2
Graphie rare	cch	<i>saccharine</i>	C

Fig.3: exemple de hiérarchisation de la relation phonème-graphème pour un phonème consonantique

Alors qu'en matière de graphèmes, plus on "monte" dans les niveaux, moins il y a de nouveaux graphèmes, c'est l'inverse pour l'homophonie lexicale: 214 entrées dès A1, 556 à compter de A2, 1290 à partir de B1, et 2097 en B2. Autrement dit, 10% des entrées-cibles recensées dans B2 doivent être reconnues dès A1, 26% (+16%) en A2, 61% (+35%) en B1, B2 ajoutant les 39% restantes. A cet aspect quantitatif s'ajoute une perspective qualitative: pour chaque entrée, le nombre de termes mentionnés augmente, en relevant les phénomènes d'homophonie des niveaux ultérieurs: *sol* est évoqué dès A1, dans la mesure où il y a une confusion possible et à venir avec *sole* et *saule*, qui n'apparaissent en tant que tels qu'en B2; on va trouver *lait*, *les* (*laid*) en A1 et A2, et *lait*, *les*, *laid* (*lai*, *laie*, *lais*, *lei*, *lé*, *lès*, *lez*) en B1 et B2. Que ces derniers termes relèvent un jour de C1, C2, ou même C2+, ne change rien au fait que l'importance de la discrimination graphique *lait*, *les*, *laid* est corrélée à l'ampleur de l'homophonie sous-jacente, qui s'étend en outre à la morphologie liée (*l'es-t*, *l'ai-e-s-nt*).

Celle-ci dépend de l'orthographe grammaticale, qui relève tout d'abord cette homophonie des mots grammaticaux, seuls ou en séquence. Comme pour les graphèmes, l'essentiel apparaît dès A1 (23 entrées, contre 34 en B2), mais, comme pour l'homophonie lexicale, il faut intégrer une perspective qualitative: dans la séquence *les* / *l'est* / *l'ai* / *l'est* / *l'es*, *l'es* est signalé en A1 et intégré à compter de B1, en même temps que l'ensemble des morphèmes de personne 2. Compte tenu de la spécificité du français, on a ensuite donné la liste des morphèmes grammaticaux invariables (45 en A1, contre 79 en B2). A compter de B2 apparaissent 17 homonymes à double genre (*un/une mode*). Pour le genre et le nombre, les procédés les plus courants (*-e* féminin, *-s/-x* pluriel) apparaissent dès A1, et les réalisations complexes à compter de B1 (*magnifique/public*, *habile/subtil*, *travail/portail*, *canal/bal*, pluriel des adjectifs ou des noms composés...⁷), en fonction de la présence des items dans les "fonctions générales et spécifiques".

7

On aura compris que ce sont les variations de réalisation qui sont ici visées, que cela concerne le singulier (**magnific-magnifique public-publique*, **habil-habile*, *subtile-subtil...*) ou le pluriel (*travails, travaux*, **portaux-portails*, **canaux-canaux*, *bals-baux...*).

Pour la morphologie verbale, l'ensemble du système (cf. Figure 4) est présent dès A1. Ce qui change véritablement, c'est l'introduction des temps verbaux. A1 et A2 ne comportent que le présent, qui autorise l'accès au passé (passé composé) comme au futur (futur périphrastique). Les bases de l'imparfait, du subjonctif présent et du futur simple apparaissent à compter de B1. L'introduction de la base du passé simple et du subjonctif imparfait est repoussée aux niveaux C, du fait notamment qu'elle a quasiment disparu de l'oral. Morphologiquement, les temps composés sont considérés comme n'ayant pas d'existence propre, c'est-à-dire qu'ils sont envisagés comme des formes simples suivies d'un participe passé, forme verbale adjective⁸. Dès A1, ceux-ci sont présentés comme partie intégrante des formes finales en /e/ ou /ɛ/, en /i/ et en /y/, c'est-à-dire en fonction de leurs réalisations orales, en concurrence avec un ensemble circonscrit d'autres formes verbales. Nous avons procédé de la même manière avec les désinences du présent des verbes en /ɛ̃/ (*craindre, peindre, tenir...*), en /õ/ (*fondre, rompre...*), en /u/ (*coudre, bouillir...*), ou en /o/-/ø/ (*valoir, vouloir...*), en les introduisant en fonction de l'apparition des verbes en questions dans les "fonctions générales et spécifiques". Les particularités orthographiques des formes verbales mentionnent dès A1 la notation des sons /s/, /k/, /g/ devant et /a/ et /o/ (*ç/ge/qu/gu + a/o*), le *e* muet dans les formes en voyelle + *-er* (*il crie, il sue...*, mais *il conclut*) et l'impératif personne 2. Nous introduisons les particularités relatives à l'imparfait (*criions, gagniez...*), celles des participes passés en /y/ (*dû*) et des verbes en *-dre* à compter de B1, en réservant les formes en *-eler/-eter* ou les verbes comme *acquérir* au niveau B2⁹.

⁸ L'accord du participe passé a d'ailleurs intérêt à être traité à la manière de Wilmet (1999) ou de Cavanna (1989), c'est-à-dire en considérant le participe comme un adjectif qui s'accorde lorsque le substantif support est connu, c'est-à-dire "placé avant", quel que soit l'"auxiliaire".

⁹ Suivant la même logique, les formes en /y/ étant introduites dès A1, nous introduisons en B1 les formes spéciales des participes passés en /y/. *Idem* pour les verbes en *-dre* (plus haut sont mentionnés en A1 des verbes comme *craindre, fondre, coudre*).

Nombre de bases	Type	Exemples
Une base	1B-type1	<i>courir</i> et composés
	1B-type2	<i>cueillir</i> , verbes en <i>-ure</i> , <i>-uer</i> , <i>-ouer</i> (<i>conclure</i> , <i>jouer</i> , <i>suer</i>)
Deux bases	2B-type1	verbes en <i>-er</i> (<i>parler</i>) et en <i>-ir</i> (<i>ouvrir</i> , <i>offrir</i>)
	2B-type2	Verbes en <i>e/ê+consonne+er</i> comme <i>jeter</i> , <i>lever</i>
	2B-type3	Verbes comme <i>payer</i> et <i>croire</i> , <i>prévoir</i> ...
	2B-type4	Verbes en <i>-ier</i> , comme <i>prier</i> , ainsi que <i>rire</i> et ses composés
	2B-type5	Verbes <i>battre</i> , <i>mettre</i> , <i>pondre</i> <i>vivre</i> ...
	2B-type6	Verbes <i>finir</i> , <i>lire</i> , <i>plaire</i> ...
Trois bases	3B-type1	Verbes <i>partir</i> , <i>dormir</i> , <i>sentir</i> <i>servir</i> ...
	3B-type2	Verbes <i>envoyer</i> , <i>voir</i> , <i>mourir</i> ...
	3B-type3	Verbes <i>devoir</i> , <i>recevoir</i> , <i>apercevoir</i> <i>boire</i>
Quatre bases	4B	Verbes <i>venir</i> , <i>tenir</i> , <i>prendre</i>
Cinq bases	5B	Verbes <i>pouvoir</i> et <i>devoir</i>
Irréguliers (hors du système des bases)	>5B	Verbes <i>avoir</i> , <i>être</i> , <i>aller</i> , <i>faire</i> , <i>savoir</i>

Fig.4: répartition des radicaux verbaux par bases

4. Matière graphique et niveaux C

Avec les niveaux C, dont les référentiels n'existent pas actuellement, la question orthographique prend en français une tournure particulière. Alors que les descripteurs du CECR disponibles sont suffisants pour les niveaux A et B, ils ne conviennent plus au-delà. Affirmer par exemple en C1 que "*La mise en page, les paragraphes et la ponctuation sont logiques et facilitants. L'orthographe est exacte à l'exception de quelques lapsus*" (voir annexe 1), et en C2 que "*Les écrits sont sans faute d'orthographe*" fonctionne peut-être avec l'espagnol ou l'italien, mais pas avec le français. D'une part, c'est utopique et, d'autre part, l'importance accordée à la maîtrise graphique est telle en français qu'il faut que les descripteurs en rendent compte. On peut d'ailleurs se demander s'il ne faudrait pas reprendre les mêmes descripteurs que pour le FLM qui, alors qu'ils divergent pour A et B, convergeraient pour C.

Cela donnerait pour C1: "*Connaît l'ensemble des exceptions et subtilités relatives à l'orthographe, même s'il a occasionnellement besoin de temps et de manuels spécialisés pour les mettre en pratique*" (voir annexe 2), et pour C2: "*L'orthographe est totalement maîtrisée sans outils spécialisés. Est capable de rectifier sans hésitation un texte éventuellement fautif*". On aboutit à une double contradiction: d'une part, la difficulté d'une parfaite maîtrise de l'orthographe, pour les francophones comme pour les

apprenants étrangers, associée à un discours peu réaliste sur la question; d'autre part, les limites du CECR dès lors qu'on touche aux niveaux C.

En effet, les descripteurs ci-dessus ne conviennent pas à bon nombre de locuteurs et on est incapable de définir de façon réaliste ce qui est exigible des secrétaires, des artisans, des cadres, des médecins... et des élèves. Alors qu'il existe bien évidemment des compétences de niveaux C relatives à l'écriture, ne serait-ce qu'en termes de lexique et d'aptitudes discursives, il ne devrait plus y en avoir en orthographe, ou plutôt l'orthographe ne devrait pas devenir une compétence supplémentaire discriminante. Avec le français c'est tout le contraire qui se produit, et il demeure au-delà de B2 un nombre conséquent de difficultés qui peuvent amener à concevoir un niveau C2+. Les rectifications de 1990, concernant d'ailleurs pour l'essentiel, outre l'accent circonflexe sur le *i*, des phénomènes qui relèvent des niveaux C. Qui plus est, la sacralisation de l'orthographe donne une importance particulière à cette compétence, et les efforts consacrés à surmonter les obstacles justifient en partie l'attachement de certains enseignants de FLE vis-à-vis d'une compétence somme toute secondaire au regard du CECR.

Il n'est pas anodin que l'élaboration des niveaux de référence par langue s'arrête aujourd'hui au niveau B2, ce qui ferait des niveaux C une finalité toute théorique. Le volume des connaissances à décrire devient abyssal; deux niveaux ne suffisent plus et on voit rapidement émerger un C2+, finalement destiné à héberger une hypothétique perfection; on est face à quelque chose qui ne peut plus être homogène. Tel apprenant-utilisateur développera des compétences en fonction de ses besoins qui pourront être importants en compréhension orale et modestes en expression écrite. Tel apprenant-utilisateur atteindra un C2+ dans l'ontologie restreinte qui est la sienne, en demeurant bien en deçà pour tout le reste. Tel apprenant-utilisateur ne conservera ses compétences que le temps qu'il les pratique au meilleur niveau, et les verra évoluer au même rythme que ses besoins.

Comme le dit Riba (2010: 425-426), "la vraie compétence acquise du locuteur de niveau C devrait être la possibilité de comprendre l'autre sans se diluer en lui". Une vraie compétence graphique de niveau C consisterait en somme, au-delà d'une hypothétique absence totale de "fautes", en une conscience de leur valeur, une aptitude à les reprendre, voire à les commenter. Une compétence de niveau C est une compétence réflexive, qui amènerait moins à ne jamais se tromper qu'à savoir éviter les "fautes", en fonction de leur nature comme des contextes, quitte à opérer les vérifications nécessaires et à solliciter de l'aide. Une compétence de niveau C consisterait également à disposer des outils, y compris grammaticaux, propres à ne pas réitérer ses éventuelles erreurs, à faire de celles-ci un processus vertueux, pouvant conduire à restructurer ses connaissances en fonction des obstacles rencontrés. Une compétence de niveau C

consisterait enfin à ne plus faire de l'orthographe un obstacle à une utilisation large de l'ensemble de ses connaissances, à ne plus éviter l'emploi d'un morphème lexical ou grammatical pour la seule raison qu'on a un doute sur une double lettre ou sur un accord excessivement sophistiqué.

Bibliographie

- Anis, J., Puech, C. & Chiss, J.L. (1988): L'écriture, théories et descriptions. Bruxelles (De Boeck).
- Beacco, J.C. (2007): L'approche par compétences dans l'enseignement des langues. Paris (Didier).
- Biedermann-Pasques, L. & Jecic, F. (dir) (2006): Les rectifications orthographiques de 1990: analyses des pratiques réelles (Belgique, France, Québec, Suisse, 2002-2004). Coll. Les Cahiers de l'Observatoire des pratiques linguistiques, No1. Orléans (Presses Universitaires d'Orléans).
- Brissaud, C., Jaffré, J.P., & Pellat, J.C. (dir) (2008): Nouvelles recherches en orthographe. Limoges (Lambert-Lucas).
- Catach, N. (1991): L'orthographe française. Paris (Nathan).
- Cavanna, F. (1989): Mignonne allons voir si la rose. Paris (Belfond).
- Conseil de l'Europe (2005): Cadre européen commun de référence pour les langues, Apprendre Enseigner, Évaluer, 2005. Paris (Didier).
- Fayol, M. & Jaffré, J.P. (2008): Orthographier. Paris (PUF).
- Gak, V. G. (1976): L'orthographe du français. Paris (SELAF, réédition Peeters 2000).
- Jaffré, J.P. & Pellat, J.C. (2008): Sémiographie et orthographe: le cas du français. In: Brissaut, Jaffré & Pellat 2008 (dir), 9-30.
- Luzzati, D. (2004): Chapitre 8: "la matière graphique". In: Niveau B2 pour le français, un référentiel, 289-326, Beacco, J.C., Bouquet, S., Porquier R. (dir). Paris (Didier).
- Luzzati, D. (2007): Chapitre 8: "la matière graphique". In: Niveau A1 pour le français, un référentiel, 139-155, Beacco, J.C, Porquier R. (dir). Paris (Didier)
- Luzzati, D. (2008): Chapitre 8: "la matière graphique". In: Niveau A2 pour le français, un référentiel, 139-155, Beacco, J.C, Porquier R. (dir). Paris (Didier).
- Luzzati, D. (2011): Chapitre 8: "la matière graphique". In: Niveau B1 pour le français: un référentiel, Beacco, J.C., Blin, B., Houles, E., Lepage, S., Riba, P. (dir). Paris, (Didier).
- Luzzati, D. (2010): Le français et son orthographe. Paris (Didier).
- Pinchon, J. & Couté, B. (1980): Le système verbal du français: descriptions et applications pédagogiques. Paris (Nathan).
- Riba, P., (2010): La spécification et la certification pour le français des niveaux C1 et C2 du Cadre européen commun de référence pour les langues Thèse, Paris 3.
- Wilmet, M. (1999): Le participe passé autrement. Paris-Bruxelles (Duculot).

Annexe 1: descripteurs du CECR pour la compétence orthographique générale, Conseil de l'Europe (2005: 92-93)

Déclinaison par niveaux de la "Maîtrise de l'orthographe" (p. 93)	
A1	Peut copier de courtes expressions et des mots familiers, par exemple des signaux ou consignes simples, le nom des objets quotidiens, le nom des magasins et un ensemble d'expressions utilisées régulièrement. Peut épeler son adresse, sa nationalité et d'autres informations personnelles de ce type.
A2	Peut copier de courtes expressions sur des sujets courants, par exemple les indications pour aller quelque part. Peut écrire avec une relative exactitude phonétique (mais pas forcément orthographique) des mots courts qui appartiennent à son vocabulaire oral.
B1	Peut produire un écrit suivi généralement compréhensible tout du long. L'orthographe, la ponctuation et la mise en page sont assez justes pour être suivies facilement le plus souvent.
B2	Peut produire un écrit suivi, clair et intelligible qui suive les règles d'usage de la mise en page et de l'organisation. L'orthographe et la ponctuation sont relativement exacts mais peuvent subir l'influence de la langue maternelle.
C1	La mise en page, les paragraphes et la ponctuation sont logiques et facilitants. L'orthographe est exacte à l'exception de quelques lapsus.
C2	Les écrits sont sans faute d'orthographe.

Annexe 2: descripteurs dérivés du CECR pour une compétence orthographique en français langue maternelle, Luzzati (2010: 239-240)

Déclinaison par niveaux d'une Maîtrise de l'orthographe dans une perspective FLM	
A1	Maîtrise l'essentiel des signes du code écrit. Peut copier un court texte presque sans erreur. Distribue correctement les blancs. Sait orthographier l'essentiel de son vocabulaire courant. Peut transcrire l'essentiel des phénomènes d'accord, dès lors qu'ils sont audibles. Peut produire un texte court parfaitement compréhensible.
A2	Connait l'ensemble des signes du code écrit. Copie sans erreur. Maîtrise parfaitement la distribution des blancs. Orthographie presque sans erreur l'essentiel du vocabulaire courant, même s'il ne le pratique pas régulièrement. Respecte globalement les accords, même ceux qui ne sont pas audibles. Peut produire un texte court pratiquement sans erreur et un texte relativement long parfaitement compréhensible.
B1	Peut produire un écrit suivi, clair et intelligible qui suive les règles d'usage de la mise en page, de l'organisation et de l'orthographe, même s'il aborde des sujets nouveaux. Peut identifier dans ses textes l'essentiel des mots susceptibles de comporter des fautes et parvient à en corriger une bonne part, avec les outils à sa disposition.
B2	Peut produire un texte long et structuré qui ne comporte plus que quelques erreurs d'orthographe résiduelles et non systématiques. Maîtrise la relecture d'un texte fautif et parvient à le rectifier sans mal, à l'aide des outils à sa disposition.
C1	Parvient à varier l'approche discursive et à produire des textes dans une langue soutenue. <i>Connait l'ensemble des exceptions et subtilités relatives à l'orthographe, même s'il a occasionnellement besoin de temps et de manuels spécialisés pour les mettre en pratique.</i>
C2	Maîtrise parfaitement les variations discursives et peut utiliser une langue recherchée autant que de besoin. <i>L'orthographe est totalement maîtrisée sans outils spécialisés. Est capable de rectifier sans hésitation un texte éventuellement fautif.</i>

Les étudiants natifs et allophones face à l'orthographe française: le cas des homonymes

Martha MAKASSIKIS & Jean-Christophe PELLAT

UR LILPA (Linguistique, langues, parole), Université de Strasbourg

As a mixed system, French spelling has both significant phonographic and semiographic characteristics. Due to the impact of spelling structure on learning and acquisition strategies, one must take into account these specificities when elaborating teaching tools. Nonetheless, most FFL spelling textbooks focus on the phonographic dimension and give little weight to semiography.

The survey on homonyms conducted among native and non native speakers of French, studying at the University of Strasbourg, unveiled the better results of the latter in writing logograms during a written examination. However, in gap filling exercises, their success rate is significantly lower than that of natives, due to difficulties encountered in vocabulary. Both groups relied heavily on meaning to write the logograms they heard. This fact shows that the semiographic dimension gains importance when phonography is powerless to help the writers.

Qui compare les manuels d'orthographe existants destinés à un public endophone (*Le Bled* (1946) et *O.R.T.H.* (1979), par exemple) et allophone (*Orthographe progressive du français* (2004) et *Orthographe 450 nouveaux exercices* (2003)) est frappé par les choix des auteurs. On note en effet un choix commun, l'étude des morphogrammes, qui semble dicté par la structure de l'orthographe française, et un choix spécifique au public visé, les logogrammes en français langue maternelle (FLM), les phonogrammes en français langue étrangère (FLE). L'objectif de cette étude est de vérifier, au moyen de deux expériences, si les zones orthographiques ciblées par les concepteurs de manuels en fonction du public sont fondées et de mettre à l'épreuve l'idée selon laquelle les logogrammes ("figures de mots"), permettant de distinguer les homophones, sont moins aisés à acquérir chez les étudiants natifs.

Après une mise au point sur les enjeux de l'apprentissage de l'orthographe française et un aperçu du traitement de l'orthographe dans quelques manuels de FLE, nous exposerons les résultats d'une double enquête: une analyse de productions écrites et un compte rendu d'entretiens sur l'orthographe et les stratégies de mémorisation des homonymes en français.¹

¹ Le présent texte applique les Rectifications orthographiques de 1990.

1. Les enjeux de l'apprentissage de l'orthographe française

1.1 *L'incidence de la structure d'une orthographe sur les apprentissages*

Diverses observations tendent à montrer que la structure des orthographe a une incidence sur les apprentissages et les stratégies d'acquisition.

Dorit Ravid (2001: 1) constate que, dans l'acquisition de l'orthographe, "children are early on sensitive to the typological imperatives of their language"². Ils construisent en effet des stratégies parfaitement adaptées aux traits typologiques de la langue dont ils doivent apprendre l'orthographe. Pour acquérir la morphologie de l'hébreu, les enfants doivent ainsi distinguer *root*, *pattern* et *stem-and-suffixe structure*³. La connaissance orthographique se subdivise alors en quatre aspects majeurs. Il s'agit d'abord de "mapping phonology onto graphemic segments"⁴ (*idem*: 11), avec les problèmes de notation qui en découlent, consonantiques et surtout vocaliques. Viennent ensuite "the conventions of the orthographic system"⁵ (*idem*: 11) et les difficultés qu'entraînent l'absence de notation vocalique et la démarcation des mots. Le troisième aspect est à mettre en relation avec les "morphological regularities in the spelling system"⁶ (*idem*: 12). Les "two morphological classes which are reflected in the orthography – root letters and function letters"⁷ (*idem*: 20) constituent une aide à l'acquisition de l'orthographe, mais "homophonous root letters are clearly more difficult to spell correctly than homophonous function letters"⁸ (*idem*: 24). Sur ce point, et pour mieux mettre en évidence les spécificités de l'acquisition en hébreu, D. Ravid propose une comparaison avec l'orthographe du néerlandais. Il s'avère, en effet, que contrairement aux enfants hébreux, les "Dutch-speaking children"⁹ n'accordent guère d'intérêt à la morphologie, "because they do not need morphological operations in order to construct words"¹⁰ (*idem*: 26).

² "les enfants sont sensibles tôt aux impératifs typologiques de leur langue" (cette traduction et celles qui suivent sont de notre initiative)

³ "la racine, le patron et la structure radical-et-suffixe"

⁴ "faire correspondre la phonologie aux segments graphémiques"

⁵ "les conventions du système orthographique"

⁶ "régularités morphologiques du système orthographique"

⁷ "deux classes morphologiques qui sont reflétées dans l'orthographe – lettres radicales et lettres fonctionnelles –"

⁸ "les lettres radicales homophones sont manifestement plus difficiles à bien orthographier que les lettres fonctionnelles homophones"

⁹ "enfants néerlandophones"

¹⁰ "parce qu'ils n'ont pas besoin des opérations morphologiques pour construire les mots"

1.2 *La structure des systèmes d'écriture et de l'orthographe française*

Tous les systèmes d'écriture suivent deux principes directeurs: la représentation du sens linguistique (*principe sémiographique*) et la représentation des unités sonores, phonèmes ou syllabes (*principe phonographique*). La linguistique structurale, suivant Ferdinand de Saussure, a privilégié le principe phonographique, en donnant comme modèle idéal de l'orthographe la transcription phonétique (position encore défendue par V. Gak en 2001). Ce faisant, elle a relégué au second plan, voire ignoré, le principe sémiographique, pourtant essentiel, puisque l'écriture en tant que moyen de communication doit donner accès au sens. En fait, les systèmes d'écriture reposent sur un compromis entre les principes phonographique et sémiographique, qui résulte de l'histoire particulière de la constitution de chaque système. On a évalué le degré de transparence phonographique des orthographe utilisant les caractères latins, en opposant le finnois, l'espagnol et le hongrois, les plus transparentes, à l'anglais, au français et au danois, les plus opaques. D'un autre point de vue, on peut distinguer les orthographe où la part de la sémiographie est importante (anglais, chinois, français, japonais) et celles où elle joue un rôle mineur (le turc par exemple). Bref, on peut dire que pratiquement tous les systèmes d'écritures sont mixtes, combinant selon des dosages différents la sémiographie et la phonographie.

L'orthographe française constitue elle aussi un *système mixte*, un *plurisystème* ou "système de systèmes" selon N. Catach (1980). Si l'on peut observer une forte régularité (à 96%) dans les correspondances graphèmes-phonèmes (règles de prononciation), les correspondances phonèmes-graphèmes (règles de transcription) sont plus irrégulières (71%, selon Fayol & Jaffré 2008: 115). Cela tient à l'importance de la dimension sémiographique de l'orthographe française: les lettres dites muettes jouent un rôle à des niveaux linguistiques supérieurs (lexique et grammaire) et les phonogrammes eux-mêmes sont utilisés à des fins sémiographiques (Jaffré & Pellat, 2008). Selon les termes de N. Catach (1980), les morphogrammes grammaticaux, qui ne sont pas toujours prononcés, jouent un rôle essentiel dans la morphologie grammaticale, comme le -s des pluriels nominaux (*chants*) ou de la 2^e personne du singulier des verbes (*tu chantes*) ou le -e du féminin (*Emilie jolie*). Les logogrammes (ou figures de mots) permettent de distinguer la plupart des homophones lexicaux ou grammaticaux du français, comme *cygnes* et *signes*, *son* et *sont*.

L'apprentissage de l'orthographe est conditionné par cette structuration complexe. En français langue maternelle, on commence à construire un modèle développemental sur certaines zones orthographiques, comme le font C. Brissaud *et alii* (2006) sur l'acquisition des formes verbales en /E/ de

6 à 15 ans: au CP, le premier traitement des formes en /E/ est phonographique (choix des graphèmes simples *e* et *é*), puis le traitement morphographique est introduit progressivement, à commencer par le choix du morphogramme *-er* dès la 2^e année du primaire. L'écriture des homophones constitue un autre problème majeur des scripteurs natifs, comme nous le verrons ci-après.

1.3 *Un exemple de structuration inadéquate du domaine orthographique dans un manuel de FLE*

Pour bien apprendre l'orthographe française, il importe donc de bien comprendre sa structure. Or, il se trouve que souvent la structure de l'orthographe française n'apparaît pas de manière claire dans les manuels qui lui sont consacrés. L'*Orthographe progressive du français, niveau débutant* d'I. Chollet et de J.-M. Robert (2004) (désormais *OPF*) en témoigne.

Ce manuel complémentaire, très utilisé dans les classes de FLE, donne la priorité aux correspondances phonographiques. Comme les apprenants du FLE rencontrent au départ des difficultés avec la prononciation du français, beaucoup de manuels de FLE subordonnent l'orthographe à la prononciation, insistant en particulier, pour la lecture, sur la distinction entre les lettres qui se prononcent et celles qui ne se prononcent pas. Dans cette intention certes louable, on ne donne guère d'indications sur les règles de transcription qui commandent le choix des graphèmes dans l'écriture du français et l'on fait souvent l'impasse sur le niveau sémiographique.

Les deux leçons choisies de l'*OPF* (2004: 8-9, 104-105) (*annexes 1a-1d*) sont le chapitre 2 "*i* ou *î* (*y*)" et le chapitre 40 "consonnes finales non prononcées (1): le *s*". Elles se découpent en deux parties: après un petit encadré contenant deux phrases illustrant les lettres à étudier, une page de leçon donne un inventaire d'observations de l'usage des lettres concernées, suivie d'exercices d'application sur la page en regard; la démarche est nettement déductive, comme on la connaît en FLM avec le Bled. Mais, contrairement à celui-ci, *OPF* formule très peu de règles ou de préceptes, se contentant le plus souvent d'aligner des observations de l'usage illustrées par divers exemples (avec des renvois à d'autres leçons), ce qui est une pratique courante des manuels de FLE.

Ces leçons comportent trois défauts principaux:

1) En ce qui concerne la lettre *i*, les correspondances avec l'oral ne sont pas hiérarchisées. Or, les trois graphies présentées n'ont pas la même importance (D. Luzzati, 2010: 24). Alors que *i* et *y* sont prioritaires au niveau débutant (niveau A du *Cadre Européen Commun de Référence pour les Langues - CECRL*), *î* peut être étudié plus tard (niveau B du *CECRL*). En

outre, on pourrait alléger le travail des étudiants en tenant compte des Rectifications de 1990: dans tous les exemples (*dîner, île, chaîne, plaît, connaître, boîte*), l'accent circonflexe peut être supprimé, ce qui rend inutile cette partie consacrée à *î*.¹¹

2) La présentation des usages de *i* et les exercices mêlent deux approches inverses des rapports phonèmes-graphèmes: le rapport de transcription (différentes graphies du phonème [i]) est compliqué par la présentation de la prononciation de digrammes qui correspondent à des phonèmes différents (*ai, oi, in, aî, oî*), ce qui brouille la perception de la prononciation du phonème directement associé au graphème étudié.

3) La dimension sémiographique est subordonnée aux correspondances phonographiques:

- Pour la lettre *i*, on présente pêle-mêle une liste de consonnes et de voyelles finales: "A la fin d'un mot, la lettre *i* est parfois suivie de *e, l, s, t, x* ou *y* qui ne se prononcent pas". Aucune explication n'est donnée sur le rôle de ces lettres, qui viennent en quelque sorte "habiller" la lettre *i* dans les mots (naguère, la méthode du Sablier (1976) présentait les "costumes" d'une lettre). On y trouve des morphogrammes lexicaux (*lit, nuit, gris*), des éléments de suffixes (*boulangerie*), des lettres à valeur logogrammique (*prix*), etc.

- Pour la lettre *s*, le titre met en valeur la subordination à l'oral (*Les consonnes finales non prononcées*), donnant une présentation négative de ses emplois. Mais on commence par la règle générale du pluriel: "Au pluriel, les noms et les adjectifs prennent généralement un *-s* muet". L'essentiel de la leçon est consacré à la problématique de la prononciation de *-s* final (ou des consonnes qui le précèdent: *poids, longtemps*), avec quelques compléments d'information éparpillés sur les noms et adjectifs terminés par *-s* (*héros, précis*), les articles, adjectifs possessifs et pronoms terminés par *-s*, et divers mots invariables (*alors, après, ...*), non concernés par le pluriel. La leçon s'achève sur le *-s* final muet "dans les conjugaisons", sans indication de la personne concernée (*tu finis, nous finissons*). L'absence de traitement spécifique des deux valeurs grammaticales principales de *-s* (pluriels nominaux et 2^e personne du singulier des verbes), associée au manque d'explications morphologiques, n'apporte aucune aide aux apprenants pour savoir utiliser le *-s* à bon escient.

Quant aux exercices suivant les leçons, ils portent sur l'orthographe de façon ciblée (exercices à trous sur la ou les lettres concernées) ou plus globale (écrire un nom correspondant à un dessin). Pour *s*, on relève un exercice classique de mise au pluriel d'un groupe nominal (Ex.: *un mot précis* → *des mots précis*) et un exercice où il faut barrer le *-s* "quand il ne

¹¹ Une preuve que les Rectifications simplifient l'apprentissage de l'orthographe en FLE.

se prononce pas". Pour les deux lettres, un exercice de remise en ordre des lettres d'un mot (anagramme) semble difficile pour les apprenants (*un hôpital ou une ncqliieu*: la réponse attendue est *clinique*). En résumé, à part les anagrammes, les exercices confirment la priorité donnée à la question de la prononciation des lettres.

1.4 Une approche linguistique de l'orthographe française

Le *Manuel d'orthographe pour le français contemporain* (Université de Neuchâtel, 2010) suit la répartition classique en deux parties – I. Orthographe grammaticale; II. Orthographe d'usage – pour mettre en évidence les dimensions morphogrammique (grammaticale) et phonogrammique de l'orthographe française. À travers leur choix d'entamer leur manuel par l'étude de l'orthographe grammaticale, les auteurs montrent qu'ils sont conscients de la dimension sémiographique très importante de l'orthographe française au niveau B du CECRL.

Les cours proposés sont détaillés et rigoureux (*cf.* la présentation de tous les cas de figure pour le pluriel des noms, p.3-5, ou la schématisation pour l'accord du participe passé, p.58) (*annexes 2a-2b*). Les exemples et le vocabulaire actualisés sont issus de listes du logiciel *Vocabprofile* et de "Corpus français" de l'Université de Leipzig. Les formes erronées dans les exercices de cacographie semblent issues de copies authentiques d'étudiants (*cf.* **model*, **aid*, **intèrne*, p.103) (*annexe 2c*). Tous ces éléments font de ce manuel un outil scientifiquement fiable, langagièrement authentique, particulièrement adapté à l'étude du français contemporain écrit.

On pourrait seulement regretter la légère différence d'approche choisie pour les deux parties. En effet, dans la première partie, centrée sur les morphogrammes, la dimension phonographique est ponctuellement convoquée pour renseigner les apprenants sur les changements de prononciation (*cf.* *œuf* [œf] / *œufs* [ø], p.4), alors que dans la seconde partie, centrée sur les phonogrammes, la dimension morphogrammique (lexicale et grammaticale) n'est pas localement exploitée pour justifier certaines consonnes finales muettes (par exemple, *un hasard*, *un art*, p.83) ou certaines formes homophoniques (par exemple, *tant* / *temps* / *t'en*, p.77). Une suggestion d'amélioration consisterait à introduire quelques explications d'ordre sémiographique dans la partie consacrée à l'orthographe d'usage, fondée essentiellement sur la phonographie, un peu à l'image (inversée) de la partie consacrée à l'orthographe grammaticale, majoritairement fondée sur la sémiographie, et contenant quelques remarques phonographiques.

L'étude de manuels d'orthographe a révélé la conscience des concepteurs d'avoir affaire à une orthographe articulée autour de deux axes:

phonographique et sémiographique. Mais la manière d'approcher ce second axe laisse à désirer. Tantôt dissimulée derrière une vitrine phonographique (*Orthographe progressive du français*), tantôt détachée de la phonographie (*Manuel d'orthographe pour le français contemporain*), la sémiographie montre aux apprenants l'un ou l'autre de ses deux profils. Comment les étudiants natifs et allophones gèrent-ils en pratique cette dimension importante de l'orthographe française? C'est ce que nous allons examiner maintenant.

2. Le rapport des étudiants natifs et allophones à l'orthographe: performances et représentations

2.1 Enquêtes réalisées

Le rapport des étudiants natifs et allophones à l'orthographe française a été étudié à partir d'une analyse de copies (a) et d'une expérimentation orthographique sur les homonymes (b).

(a) Pour rendre compte des erreurs commises par les étudiants francophones et allophones, nous avons recueilli 122 copies d'examen d'étudiants de licence et de master inscrits dans les filières de droit et de lettres de l'Université de Strasbourg pendant l'année universitaire 2007-2008. La moitié des copies avaient pour auteurs des locuteurs-scripteurs allophones, l'autre moitié des locuteurs-scripteurs natifs. Les erreurs d'orthographe relevées ont été classées en erreurs à dominante phonogrammique, morphogrammique, logogrammique, étymologique et historique.

(b) L'expérimentation sur les homonymes s'est déroulée en trois étapes. Suite à quelques questions préliminaires qui nous ont permis d'esquisser le profil linguistique de l'étudiant (notamment son rapport à la langue française), nous avons proposé au candidat de compléter sous notre dictée un exercice à trous (*annexe 3a*). Chaque étudiant a ensuite été convié à justifier quelques graphies qu'il avait proposées. La dernière étape de l'expérimentation consistait en la lecture d'un dialogue extrait d'un livre de jeunesse, où l'on apprend la comparution au tribunal d'un petit garçon à cause d'une faute d'orthographe qu'il a commise dans une dictée (*annexe 3c*). Le dialogue a été lu par l'expérimentateur de manière théâtralisée afin de faciliter l'accès au sens. Les étudiants avaient sous les yeux une copie du texte lu et pouvaient voir les phrases en même temps qu'ils les entendaient. La lecture du dialogue était suivie de questions ciblées sur le texte ("résumez la situation décrite dans le dialogue", "que pensez-vous de cette situation? est-elle exagérée, reflète-t-elle en partie la réalité?"), de questions plus générales sur le rôle de l'orthographe, le rapport des scripteurs en général et des étudiants interrogés en particulier à

l'orthographe française (sont-ils en bons termes avec l'orthographe?), et de questions plus personnelles sur les secteurs de l'orthographe qui les fâchent le plus, les stratégies qu'ils mettent en place pour écrire les homonymes et les souvenirs qu'ils ont gardés de la manière dont ils ont appris l'orthographe française durant leurs premières années d'étude de la langue française. Soixante sujets ont été soumis à l'expérimentation: 30 sujets FLE et 30 FLM. Les volontaires allophones étaient tous étudiants. Cinq des candidats endophones étaient des retraités.

La visée de ces deux enquêtes était de comparer les performances orthographiques (en général, puis dans la zone logogrammique en particulier) des étudiants FLE pris dans leur globalité, dans toute leur diversité linguistique, à celles des étudiants FLM pour voir si de grandes tendances se dessinaient en fonction du rapport des étudiants au français – natifs ou non natifs.

2.2 Résultats obtenus

2.2.1 Analyse de copies: les étudiants allophones sont plus forts dans l'écriture des homonymes en situation de production.

	Erreurs à dominante phonogrammique	Erreurs à dominante morphogrammique	Erreurs à dominante logogrammique	Erreurs à dominante étymologique et historique	Total
56 étudiants FLE	1091 53%	669 33%	136 6,6%	150 7,4%	2046
56 étudiants FLM	742 43%	561 32,5%	155 9%	268 15,5%	1726

Tableau 1: Résultats comparés des copies des étudiants FLE et FLM.

Les résultats révèlent une répartition des erreurs orthographiques à peu près similaire dans les deux groupes. En FLM comme en FLE, il apparaît que les deux zones à haut risque sont les *phonogrammes* (graphèmes renvoyant à des phonèmes) et les *morphogrammes* (graphèmes renvoyant à des morphèmes). Les erreurs enregistrées sur les phonogrammes en FLM (43%) sont surprenantes dans la mesure où les étudiants natifs ne devraient pas rencontrer de difficulté particulière à transcrire les sons de leur langue qu'ils ont appris depuis leur plus tendre enfance. En examinant le type d'erreurs commises par les étudiants FLM dans cette zone, on constate qu'il s'agit essentiellement d'erreurs d'accents (**prejudice*, **différents*, **particulièrement*, **d'après*, **réservé*, **complètement*) et d'erreurs sur les consonnes doubles (**inssufisamment*, **unilaterale*, **bilaterrale*). En FLE, les erreurs phonogrammiques concernent aussi beaucoup d'erreurs d'accents (**pere*, **enoncés*, **séparement*, **premières*,

religieuses, *alinèa, *trés, *réél, *éxplications, *préscrit*), mais également des erreurs dues à une correspondance phonie-graphie non encore stabilisée (beacoup, *on vera, *facon, *rejeté, *decission, *premier partie, *le problem, *précisemment*) et à un problème de discrimination des sons du français (**fesceux* (pour *faisceau*), **enversement* (pour *inversement*), **aucune indice*).

Dans les deux groupes, les zones à risque moindre sont les logogrammes et les lettres étymologiques et historiques. Le risque d'erreur est moins important, sans doute parce que le nombre de mots touchés par le phénomène d'homonymie ou ayant gardé une trace écrite de leur histoire, de leur étymologie est plus limité. On note néanmoins une petite différence: les étudiants étrangers commettent moins d'erreurs sur les logogrammes et les lettres étymologiques et historiques que les étudiants français (6,6% *vs* 9% et 7,4% *vs* 15,5%). Comparativement, les étudiants FLE sont plus performants dans l'écriture des lettres étymologiques et historiques que leurs homologues français (2 fois moins d'erreurs). Notre hypothèse est qu'ayant une langue maternelle à laquelle ils peuvent se référer, ils auraient tendance à associer les lettres étymologiques et historiques du français, généralement inaudibles ou insuffisamment audibles, à leurs équivalents souvent audibles dans leur langue, ce qui aide à les écrire correctement. Pour les logogrammes, le passage par l'équivalent en langue maternelle ou par une autre langue connue des apprenants leur permettraient de distinguer clairement ce qui en français est identique phonétiquement.

Comparativement, il apparaît à priori que les étudiants français ont plus de difficultés avec les lettres étymologiques et historiques et les logogrammes, et les étudiants étrangers avec les phonogrammes. Signalons toutefois que la zone morphogrammique touche autant les FLM que les FLE (32,5% et 33% d'erreurs respectivement). Ce résultat s'explique notamment par le fait que la morphographie est une dimension bien développée et caractéristique du système orthographique français.

Ces résultats sont corroborés par une analyse statistique au moyen du test de χ^2 (khi deux) ou test de Pearson, qui "sert à apprécier en probabilité l'écart constaté entre une observation et un modèle théorique, quel que soit le nombre de variables" (Muller, 1992: 116).

Les étudiants FLE commettent plus d'erreurs phonogrammiques, et les étudiants FLM plus d'erreurs logogrammiques et étymologiques / historiques.

Données sélectionnées pour effectuer le test

	P	M	L	LEH
FLE	1091	669	136	150
FLM	742	561	155	268

Test de χ^2

4 fois 2 données (3 degrés de liberté).

Effectifs réels	P	M	L	LEH	Total
FLES	1091	669	136	150	2046
FLM	742	561	155	268	1726
TOTAL	1833	1230	291	418	3772

Effectifs théoriques	P	M	L	LEH	Total
FLES	994,25186	667,17391	157,84358	226,73065	2046
FLM	838,74814	562,82609	133,15642	191,26935	1726
TOTAL	1833	1230	291	418	3772

Parts de χ^2	P	M	L	LEH	Total
FLES	9,41432	0,00500	3,02288	25,96734	38,40954
FLM	11,15973	0,00592	3,58332	30,78168	45,53066
TOTAL	20,57405	0,01092	6,60620	56,74902	83,94019

$\chi^2 = 83,94019$. Probabilité < 0, 001.

Résultats

Les étudiants de FLE font nettement plus d'erreurs à dominante phonogrammique (P) et nettement moins d'erreurs à dominante logogrammique (L) et étymologique et historique (LEH) que les étudiants de FLM.

Tableau 2: Test de χ^2 effectué sur les quatre types d'erreurs orthographiques commises par les étudiants FLE et FLM: plus d'erreurs phonogrammiques chez les premiers, plus d'erreurs logogrammiques et étymologiques / historiques chez les seconds.¹²

¹² Nous tenons à remercier vivement M. Marc Hug, professeur émérite à l'Université de Strasbourg, pour son aide dans la constitution et l'exploitation de ces données statistiques.

Cette première enquête confirme les orientations des concepteurs de manuels d'orthographe. La zone morphogrammique est une zone de difficulté commune aux deux groupes. Les zones logogrammique et étymologique / historique posent plus de difficultés aux étudiants natifs, et la zone phonogrammique est plus délicate à acquérir pour les étudiants allophones. Afin d'en savoir un peu plus sur les raisons des meilleures performances des étudiants FLE dans l'écriture des homonymes, nous avons effectué une enquête complémentaire, dont nous rendons compte ci-après.

2.2.2 Expérimentation sur les homonymes: les étudiants francophones sont plus performants dans l'écriture des homonymes dans un exercice à trous.

ÉTAPE 1: ÉCRITURE D'HOMONYMES

Les résultats obtenus à l'exercice à trous apparaissent en contradiction avec ceux obtenus en situation de production écrite: les étudiants allophones qui s'étaient révélés plus performants dans l'écriture des homonymes en situation de production écrite, voient leurs performances baisser au profit des natifs dans l'exercice à trous (*annexe 3b*).

Ces performances orthographiques amoindries s'expliquent par la *fréquence de rencontre et d'utilisation* des mots à écrire. On observe en effet, dans les deux groupes d'étudiants, un faible pourcentage de réussite lorsque les mots sont rarement vus et utilisés (*chère, à l'envi*), et un fort pourcentage de réussite lorsque les mots sont fréquemment rencontrés et employés (*son, ce soir, voir*). Les étudiants allophones étant moins fréquemment exposés aux mots de la langue française et ayant moins souvent l'occasion de les employer au quotidien, leurs performances orthographiques sont inférieures à celles de leurs collègues natifs.¹³ Les performances orthographiques constatées s'expliquent aussi par le *sens* des mots à écrire: les mots mal orthographiés sont souvent ceux qui sont inconnus des sujets.

ÉTAPE 2: EXPLICATION DE GRAPHIES

Les sujets ont invoqué 12 raisons principales pour expliquer leurs graphies.

- *la grammaire*: "Les quelques fois: *quelque* s'accorde avec *fois*, parce que c'est un adjectif dans ce cas." - "*Quel qu'il soit*: lorsque le pronom *quel* est suivi de *que* et du subjonctif, il ne peut être lié." -

¹³

On peut noter au passage que les étudiants de FLM ont obtenu un pourcentage de réussite record pour l'écriture de *cygnes* (100% contre 40% seulement en FLE). Les deux groupes ont obtenu un pourcentage de réussite identique pour l'écriture de *quelques fois* (87%).

"J'ai écrit *basilic* avec -c, parce que c'est un nom masculin. D'habitude, c'est le féminin qui s'écrit avec -que."

- *le sens*: "*Près*: la proximité est mise en valeur, plutôt que le fait d'être sur le point. Et comme dans la phrase *le jour est près de disparaître*, *le jour* n'est pas un être animé, il ne peut pas être suivi de *prêt* avec un -t."
- *le son*: "*Loncre*: j'ai écrit ce que j'ai écouté [*sic*]" - "*Les seins mènent une vie acèptique*: je sais que ce n'est pas ça, mais c'est ce que j'ai entendu." - "(*coup de*) *poing*: je prononce en moi-même la consonne inaudible avec un petit accent du sud pour me souvenir de l'orthographe."
- *la manipulation d'outils linguistiques* (remplacement par un synonyme, mise au féminin, etc.): "*Les quelques fois*: j'ai pensé à *plusieurs fois*." - "*Le jour est près de disparaître: le jour est prêt*, mais je ne peux pas dire *la journée est prête*; donc c'est *près*." - "*C'est au patron que tu auras affaire*: j'ai écrit *affaire* en un mot, parce que c'est la construction *avoir affaire à qqn*, et non *avoir qqch à faire*, qui lui s'écrit en deux mots."
- *l'association avec d'autres mots français de la même famille, des mots étrangers équivalents qui aident à lever l'ambiguïté graphique ou des mots anciens (étymons)*: "*Le cours du fleuve: cours*, ce n'est jamais que le masculin de *course*." - "*Tâche* dans le sens de *mission* s'écrit avec un accent circonflexe. Je l'associe au roumain *tașka*, qui signifie *petit bagage qu'on doit porter*." - "*Ascétique* vient du grec ασκητικός."
- *la conformité au patron graphique français*: "J'ai écrit *basilique* avec -que, parce qu'avec -c, ce n'est pas très français." - "*Ascétique* me semble trop simple sans le *h*. En français, on aime bien les *h*, on aime bien faire compliqué."
- *la mémoire visuelle*: "Je connais l'orthographe du mot *basilic*, car je le vois souvent écrit sur mon lieu de travail." - "*À l'envi*: je l'ai vu dans ma tête sans -e." - "J'apprends les mots d'une manière photographique."
- *la routine*: "*Censé*: c'est comme ça que je l'écris d'habitude. C'est la première orthographe qui m'est venue à l'esprit." - "*Exigeants*: dérive du verbe *exiger*. Je ne suis pas toujours consciente pourquoi je l'écris comme ça. Ça vient spontanément." - "*Cours*: le mot m'est venu naturellement."
- *l'esthétique*: "*Appelle*: c'est plus joli avec deux *l*." - "*Chaire*: ça ne me plaît pas comme je l'ai écrit."

- *l'incertitude*: "*Prêt*: j'ai un gros doute. *Près*, c'est proche et *prêt*, il est l'heure. Je n'arrive pas à trancher." - "*Appelle*: je ne sais pas s'il y a un accent; je ne me souviens pas."
- *la certitude*: "*Coup*: je suis sûre, parce que je le sais." - "*Voir*: je connais le mot."
- *la norme absolue*: "*Cris*: parce que je sais que ça s'écrit comme ça." - "*Aire*: c'est comme ça."

De manière générale, les trois explications le plus fréquemment émises dans les deux groupes ont été le sens du mot, la grammaire et la mémorisation. En FLE, la phonétique a également été invoquée à de multiples reprises pour expliquer la graphie des mots. Cette explication serait-elle la trace d'une représentation essentiellement phonographique de l'orthographe française chez les étudiants allophones? Il semblerait que oui. Ce fait confirme l'idée que l'orthographe à apprendre (ici alphabétique) détermine la stratégie de tout scripteur (d'abord phonographique).

Cette seconde enquête permet de nuancer les résultats de l'analyse des copies. En situation de production écrite, les étudiants allophones commettent moins d'erreurs dans l'écriture des homonymes, mais ils se révèlent plus vulnérables dans l'écriture d'homonymes choisis et dictés par un référent externe. Les mots qu'ils sont alors invités à écrire ne leur sont pas toujours connus. Parfois, ils sont connus d'eux, mais non reconnus à l'oral, en raison de la prononciation française qui n'oralise pas tout ce qui est écrit. La séquence phonique opacifie l'accès au sens et complique la tâche de sélection de la graphie adéquate.

2.2.3 Représentations de l'orthographe française

ÉTAPE 3: ENTRETIEN SEMI-DIRIGÉ

Dans la dernière étape de l'expérimentation, nous souhaitons mettre au jour les stratégies employées par les deux groupes pour bien orthographier les homonymes. Nous n'avons pas trouvé de stratégies différentes en fonction du rapport des sujets au français. Les stratégies étaient dans l'ensemble les mêmes: délimitation du sens du mot à partir du contexte donné dans la phrase, attention portée à l'environnement immédiat du mot dans la phrase (ce qui précède et suit), mémorisation de l'image du mot...

Toutefois, nous avons pu faire ressortir trois *représentations* des locuteurs-scripteurs français *sur l'orthographe de leur langue* (Figure 1).

- La première représentation de l'orthographe, décelée chez deux des cinq retraités, âgés de 63 et 68 ans, consiste à penser que faire des *fautes* d'orthographe dans un écrit qu'on adresse à quelqu'un, c'est *faire injure* à cette personne, lui manquer de respect.

R1: J'offense l'autre par mes fautes d'orthographe.

- La deuxième représentation, repérée chez des apprenants âgés de 20 à 23 ans, consiste à dire que faire des *fautes* d'orthographe dans un écrit qu'on adresse à quelqu'un, c'est *lui laisser la possibilité de nous juger*, c'est se rabaisser aux yeux de cette personne.

R2: Je suis jugé par l'autre à cause de mes fautes d'orthographe.

- La troisième représentation, identifiée chez un étudiant de 23 ans, voit les problèmes d'orthographe non plus comme des fautes, mais comme des *erreurs* qui sont *la marque d'une pratique insuffisante de l'écrit*, d'un raisonnement différent de celui institué par la norme, d'une systématisation insuffisante. Commettre des erreurs d'orthographe, c'est un peu normal au vu de la complexité de l'orthographe française; c'est gênant quand on exerce un métier où l'on a besoin de maîtriser l'orthographe; autrement, ce n'est pas dérangeant. Mais cela ne doit en aucun cas être utilisé pour émettre un jugement de valeur sur autrui.

R3: Mes erreurs d'orthographe ne portent nullement atteinte à ma personne, ni à celle d'autrui. Elles me signalent que je dois poursuivre mes efforts en matière d'acquisition de l'orthographe.

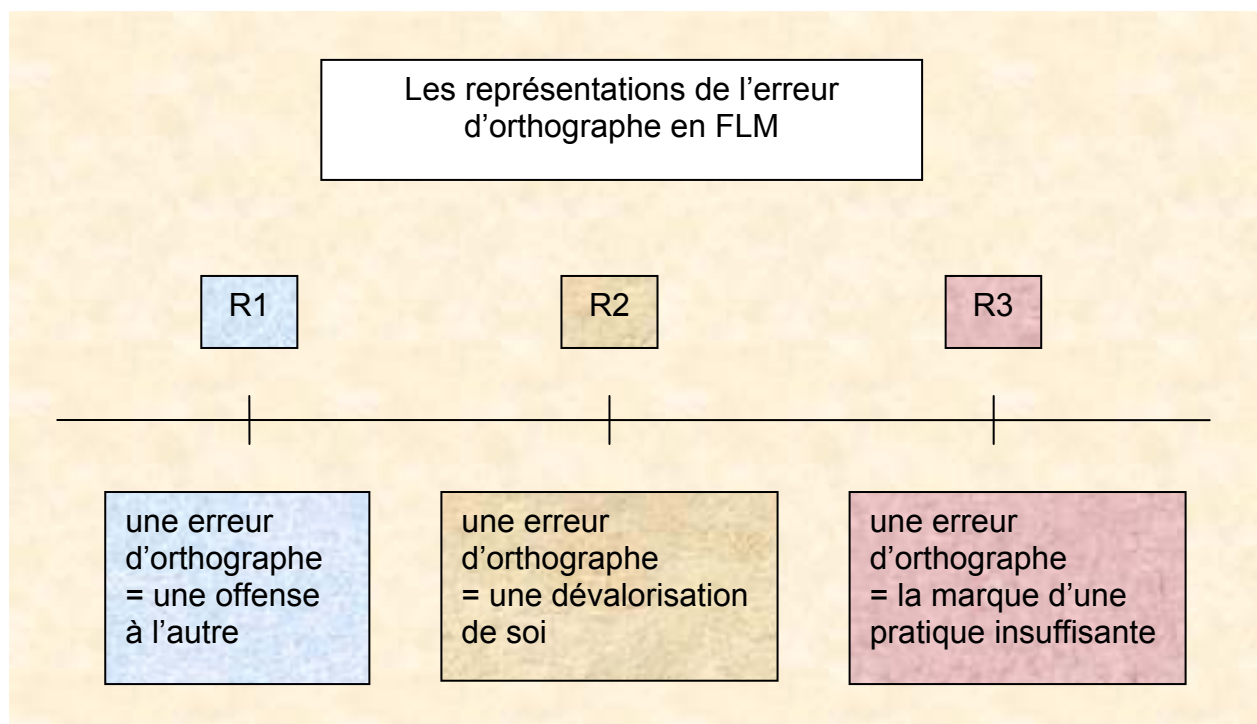


Fig.1: L'erreur d'orthographe en FLM: une histoire de représentation(s).

3. Conclusion

En conclusion, les apprenants allophones rencontrent des difficultés avec les phonogrammes français et les apprenants natifs comme les allophones gèrent mal les sémiogrammes. Toutefois, la relation n'est pas la même en fonction du sémiogramme considéré, du contexte d'écriture et du rapport de l'apprenant au français. Alors que les morphogrammes posent autant de problèmes aux natifs qu'aux allophones, les logogrammes posent plus de difficultés aux natifs en situation de production écrite et plus de difficultés aux allophones dans le contexte d'un exercice à trous. Pour contrer les pièges liés à la prononciation identique des homonymes, les apprenants des deux groupes s'appuient sur le sens du mot qu'ils ont à écrire, ainsi que sur des indices grammaticaux (syntagmatiques et paradigmatisques). Quand ces deux facteurs se révèlent inefficaces, ils recourent à la mémorisation.

La structure de l'orthographe française est complexe parce qu'elle articule, comme toutes les orthographe, deux dimensions (phono- et sémiographique), mais qui peuvent, contrairement aux autres orthographe, être combinées d'au moins quatre manières différentes: la phonographie seule (ex. *a* d'*arrivée*), la sémiographie seule (ex. *-e* d'*arrivée*), la phonographie assortie d'éléments sémiographiques (ex. *est*), la sémiographie assortie d'éléments ayant une valeur phonographique inexprimée dans ledit mot, mais qui se réalise dans les dérivés (ex. *-d* de *retard*). Les manuels devront en conséquence s'efforcer de présenter ces deux dimensions de manière à la fois séparée et combinée – la phonographie seule et en lien avec la sémiographie; la sémiographie seule et en lien avec la phonographie –, puisque l'on sait que la structure d'une orthographe détermine les apprentissages et stratégies d'acquisition. En FLE, la priorité phonogrammique déterminée par l'apprentissage de la prononciation du français ne saurait exclure la dimension sémiographique.

Enfin, les représentations sur l'orthographe française varient en fonction du public cible. Les étudiants allophones ont une représentation essentiellement phonographique de l'orthographe qui correspond à la manière dont cet objet leur est enseigné. Les étudiants endophones nourrissent une représentation plutôt normative de l'orthographe, qui est sans doute conforme à la représentation qu'on en a en France et qui se manifeste à l'occasion des grands débats linguistiques (Rectifications orthographiques de 1990, débats actuels sur la féminisation des noms de professions, etc.).

Bibliographie

Manuels d'orthographe

- Bled, E. & Bled O. (1946): Cours d'orthographe et Cours supérieur d'orthographe. Paris (Hachette éducation).
- Chollet, I. & Robert, J.-M. (2004): Orthographe progressive du français, niveau débutant. Paris (CLE International).
- Guion, J. & J. (1979-1983): Apprendre l'O.R.T.H.OGRAPHE: livre de l'élève CP, CE1, CE2, CM1, CM2, 6e et 5e, 4e, 3e, LEP. Paris (Hatier).
- Hermeline, L. (2003): Orthographe 450 nouveaux exercices – Niveau débutant. Paris (CLE International).
- Skupien Dekens, C., Kamber, A. & Dubois, M. (2010): Manuel d'orthographe pour le français contemporain. Université de Neuchâtel.
- Thimonnier, R. (1974): Code orthographique et grammatical. Verviers (Marabout).

Ouvrages de référence

- Brissaud, C. & Cogis, D. (2011): Comment enseigner l'orthographe aujourd'hui? Paris (Hatier).
- Brissaud, C., Chevrot, J.-P. & Lefrançois, P. (2006): Les formes verbales en /E/ entre 8 et 15 ans: contraintes et conflits dans la construction des savoirs sur une difficulté majeure du français. In: Langue française, 151, 74-93.
- Brissaud, C., Jaffré, J.-P. & Pellat, J.-C. (éds.) (2008): Nouvelles recherches en orthographe. Limoges (Lambert-Lucas).
- Catach, N. (1980): L'orthographe française. Paris (Nathan).
- Coulmas, F. (2003): Writing systems. An introduction to their linguistic analysis. Cambridge (Cambridge University Press).
- Cuq, J.-P. & Gruca, I. (2005): Cours de didactique du français langue étrangère et seconde. Grenoble (PUG).
- Ducard, D., Honvault, R. & Jaffré, J.-P. (1995): L'orthographe en trois dimensions. Paris (Nathan).
- Fayol, M. & Jaffré, J.-P. (2008): Orthographier. Paris (PUF).
- Gak, V. (1976): L'orthographe du français: essai de description théorique et pratique. Paris (SELAF).
- Gak, V. (2001): À propos du système graphique du français: quelques problèmes à discuter. In: Gruaz C. & Honvault R. (éds): Variations sur l'orthographe et les systèmes d'écriture, Paris (Champion), 23-34.
- Gruaz, C. (2005): Aspects du mot français. Paris (L'Harmattan).
- Harris, R. (1993): La sémiologie de l'écrit. Paris (CNRS Éditions).
- Heniqui, F. (2006): L'interprétation sémiographique des phonogrammes: les voyelles du français [o] et [ē]. Mémoire de Master réalisé sous la direction de J.-C. Pellat. Université de Strasbourg.
- Jaffré, J.-P. & Pellat, J.-C. (2008): Sémiographie et orthographes: le cas du français. In: C. Brissaud, J.-P. Jaffré & J.-C. Pellat (éds.): Nouvelles recherches en orthographe. Limoges (Lambert-Lucas), 9-30.
- Jaffré, J.-P. (2001): Écriture(s) et problèmes terminologiques. In: Colombat, B. & Savelli, M. (éds.): Métalangage et terminologie linguistique. Leuven (Peeters). Coll. Orbis/Supplementa 17, 529-543.

- Jaffré, J.-P. (2003): La morphographie du français: un cas sémiographique. In: Rééducation orthophonique, 213, 13-26.
- Luzzati, D. (2010): Le français et son orthographe. Paris (Didier).
- Muller, C. (1992): Initiation aux méthodes de la statistique linguistique. Paris (Honoré Champion).
- Pellat, J.-C. (2003): Variation et plurisystème graphique au XVIIe siècle. In: Faits de langue, 22, 139-150.
- Préfontaine, R.-R. & Cote Préfontaine, G. (1976): Le Sablier. L'apprentissage de la langue, principes et techniques. Paris (Hatier).
- Ravid, D. (2001): Learning to spell in a Semitic language: the case of Hebrew. Colloque organisé par l'Université Marc Bloch – Strasbourg 2 et le CNRS. Strasbourg University, 29 & 30 novembre 2001, communication non publiée.
- Saussure, F. (de) (2001): Cours de linguistique générale. Paris (Payot).
- Spache, G. (1940): A critical analysis of various methods of classifying spelling errors. In: Journal of Educational Psychology, vol. 31, Issue 2, 111-134.
- Vachek, J. (1973): Written Language: General Problems and Problems of English. The Hague (Mouton).

Annexe 1a: Chollet, I. & Robert, J.-M. (2004): Orthographe progressive du français, niveau débutant, chapitre 2: i ou y (î) – leçon.

2

I ou Î? (Y)

Les exercices de ce livre sont faciles.
Il y a du riz pour le dîner.

La lettre I

*C'est une idée idiote.
Une carte d'identité.*

⚠ À la fin d'un mot, la lettre **i** est parfois suivie de **e, l, s, t, x** ou **z** qui ne se prononcent pas : *boulangerie, j'ai envie, maladie, pluie, pharmacie, sortie, gentil, permis, souris, fruit, lit, nuit, gris, tant pis, prix, riz*, etc.

*Après la pluie, le beau temps.
Au lit les enfants! Bonne nuit!
Quel est le prix de ce paquet de riz?*

⚠ On trouve parfois **i** :

- dans **ai** et **ei** (cf. 9 – Ai ou Ei?) : *aimer, j'ai, je sais, mais, vrai, beige*...
- dans **oi** (cf. 11 – Oi ou Oy?) : *moi, toi, trois, voici, voiture*...
- dans **in** (cf. 15 – In ou Un) et **im** (cf. 55 – La lettre M devant M, B et P) : *inviter, matin, vin, voisin, impossible, simple*,...

Le matin, Antoine boit du lait.

La lettre Î

Dans certains mots, on écrit **î** avec un accent circonflexe : *dîner, île*, etc.

Nous allons dîner au restaurant.

⚠ On trouve parfois **î** dans **ai** (cf. 9 – Ai ou Ei?) et **oi** (cf. 11 – Oi ou Oy?) : *chaîne, s'il vous plaît, connaître, boîte*, etc.

Donnez-moi cette boîte, s'il vous plaît.

La lettre Y (cf. 13 – La lettre Y)

Quelques mots s'écrivent avec un **y** : *bicyclette, gymnastique, pyjama, stylo, type, y*, etc.

Il y a un stylo sur la table.

⚠ On trouve parfois **y** dans **ay** ou dans **oy** : *essayer, payer, pays, voyage*, etc.

*Bon voyage!
Vous payez par carte bancaire?*

Annexe 1b: Chollet, I. & Robert, J.-M. (2004): Orthographe progressive du français, niveau débutant, chapitre 2: i ou y (ï) – exercices.

I ou Î?

E X E R C I C E S

1 Remettez les lettres des mots dans l'ordre.
Ex. : Un copain ou un (mia) ami.

1. Un hôpital ou une (ncqliieu) _____.
2. La pièce de la maison où l'on prépare à manger : la (ecuiinsi) _____.
3. Si je vous offre un café, vous me dites (cermi) _____.
4. Pouvez-vous me (deri) _____ où est la gare ?
5. Il est douze heures ou (iimd) _____.
6. Mon frère n'est pas (cii) _____, il est en Espagne.

2 Complétez les mots par i, ie, il, is, it, ix ou iz.
Ex. : La pharmacie.

1. Un fru_____.	6. Minu_____.	11. Lund_____.	16. La plu_____.
2. Le pr_____.	7. Mid_____.	12. Mercred_____.	17. Une sour_____.
3. La nu_____.	8. Mon cher_____.	13. Un l_____.	18. Mon mar_____.
4. Une malad_____.	9. Gent_____.	14. Du r_____.	19. La sort_____.
5. Gr_____.	10. Tant p_____.	15. La boulanger_____.	20. Le perm_____.

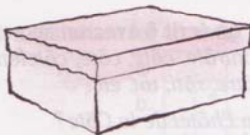
3 Complétez le dialogue par ai, ei ou oi, comme dans la première réponse.

- Elle est à toi cette v_____ture b_____ge ?
- Oui, elle est à m_____. Pourqu_____ ?
- Je l'_____me beaucoup, m_____s je la trouve trop grande.
- Tu s_____, j'_____tr_____s enfants, il me faut de la place !
- C'est vr_____.

4 Écrivez le nom de chaque objet.



Ex. : Un stylo



1. Une b_____



2. Un p_____



3. Une c_____



4. Un p_____



5. Une b_____

neuf • 9

Annexe 1c: Chollet, I. & Robert, J.-M. (2004): Orthographe progressive du français, niveau débutant, chapitre 40: les consonnes finales non prononcées (1): le *s* – leçon.

40

LES CONSONNES FINALES NON PRONONCÉES (1)

le S

Je prends un cours de français.
Il fait mauvais temps dans ce pays.

- Au pluriel, les noms et les adjectifs prennent généralement un *-s* muet : *des citrons*, *des oranges*, *des pommes rouges*, etc. (cf. 51 – Le pluriel des noms, et 52 – Le pluriel des adjectifs).

- Des mots ont un *-s* final qui ne se prononce pas.
 - Noms : *bois*, *bras*, *cas*, *cours*, *dos*, *fois*, *héros*, *jus*, *palais*, *pas*, *pays*, *repas*, *souris*, *tapis*, *velours*, etc.
 - Dans ce pays, j'ai visité un joli palais.*
 - Adjectifs : *bas*, *épais*, *français*, *gris*, *gros*, *mauvais*, *précis*, *surpris*, etc.
 - Il y a un gros nuage gris dans le ciel.*

- ⚠ Au pluriel, les noms et les adjectifs en *-s* ne changent pas : *un héros* / *des héros* ; *précis* / *précis*, etc.
 - Articles : *des*, *les*. Adjectifs possessifs : *mes*, *tes*, *ses*, *nos*, *vos*, *leurs*. Pronoms : *nous*, *vous*, *ils*, *elles*.
 - Vous connaissez mes parents ?*
(Devant une voyelle, il faut faire la liaison : *Mes amis* ; *ils aiment*.)
 - Et aussi : *alors*, *après*, *ailleurs*, *dans*, *depuis*, *jamais*, *mais*, *moins*, *plusieurs*, *près*, *sans*, *sous*, *très*, *vers*, etc. (cf. 59 et 60 – Les mots invariables [1 et 2]).

- ⚠ À la fin de certains mots, le *-s* final se prononce : *as*, *atlas*, *autobus*, *bus*, *cactus*, *ours*, etc.

- Le *-s* final muet peut être précédé d'une autre consonne muette : *-ds* et *-ps* : *poids* ; *corps*, *temps*, *longtemps*, *printemps*, etc.

- Dans les conjugaisons, le *-s* final est muet (cf. 43 à 45 – Le présent [1, 2 et 3]) : *tu as*, *tu es*, *je vais*, *tu vas*, *tu parles*, *je finis*, *tu finis*, *nous finissons*, etc.

- ⚠ Le *-s* est parfois précédé d'une autre consonne muette : *je mets*, *je prends*, *tu prends*, *je vends*, etc.

Annexe 1d: Chollet, I. & Robert, J.-M. (2004): Orthographe progressive du français, niveau débutant, chapitre 40: les consonnes finales non prononcées (1): le s – exercices.

Les consonnes finales non prononcées (1) le S

E X E R C I C E S

1 Mettez au pluriel.
Ex. : Un mot précis → *des mots précis.*

- Un étudiant français → _____.
- Un sac rouge → _____.
- Un repas fantastique → _____.
- Un palais blanc → _____.
- Un livre épais → _____.

2 Écrivez le nom correspondant à l'image.



1. _____



2. _____



3. _____



4. _____

3 Remettez les lettres des mots manquants dans l'ordre.
Ex. : – Comment allez-vous ? – (è s T r) Très bien, merci.

- Arthur est bon en mathématiques (a s m i) _____ il n'est pas bon en chimie.
- Où est mon portable ? – Il est (n s a d) _____ ton sac !
- (è A r p s) _____ la pluie, le beau temps.
- Je t'attends (i e u s d p) _____ une heure !
- Il gagne (i s m n o) _____ d'argent que sa femme.

4 Barrez le -s final quand il ne se prononce pas (attention, si un mot contient une consonne muette avant le -s final muet, rayez les deux lettres)
Ex. : Je n'ai pa~~s~~ le temp~~s~~ de vou~~s~~ parler.

- J'ai plusieurs cactus dans mon appartement.
- C'est une très bonne crème pour le corps.
- Ils sont surpris.
- Nous ne faisons jamais de fautes.
- Mon ami anglais aime les repas français.

5 Ajoutez le -s final muet quand c'est nécessaire.
Ex. : Tu fais une fête pour ton anniversaire ?

- Ce soir, nous dinon~~s~~ ensemble, tu e~~s~~ d'accord ?
- Bien sûr, je vai~~s~~ me préparer.
- Qu'est-ce que tu va~~s~~ mettre ?
- Ma nouvelle robe bleue, elle me va~~s~~ bien, n'est-ce pas ?
- Oui, très bien. Je fini~~s~~ à 19 heures et je passe~~s~~ te prendre.

cent cinq • 105

Annexe 2a: Skupien Dekens, C., Kamber, A. & Dubois, M. (2010): Manuel d'orthographe pour le français contemporain, Université de Neuchâtel, 3-7.

1. Le pluriel des noms

Les **voleurs**, originaires de divers **pays** européens, sévissaient essentiellement dans les **châteaux** et les **maisons** bourgeoises.

La phrase ci-dessus contient plusieurs **noms au pluriel** : *voleurs, pays, châteaux* et *maisons*.

- Le pluriel des noms est marqué à l'écrit par l'adjonction d'un *-s* final (*voleurs, maisons*), à moins que le nom ne se termine déjà par
- s* (*un pays* > *des pays*),
 - x* (*un prix* > *des prix*)
 - ou *-z* (*un nez* > *des nez*) au singulier.

Il existe cependant de nombreux cas particuliers.

Cas particuliers

1. Sans changement perceptible à l'oral : adjonction d'un *-x*

1.1. Les noms en *-au* (sauf : *des landaus*)

- (a) La manœuvre a été rendue possible par l'utilisation de **tuyaux** d'une longueur exceptionnelle.
- (b) Mais partout la vie continue, avec **landaus** dans les squares, bistrots bruyants et badauds mains dans les poches.

1.2. Les noms en *-eau*

- (a) Par deux fois, sa maison a été envahie par les **eaux**.
- (b) Les **bateaux** récents ne sont vraiment pas adaptés à notre région.

1.3. Les noms en *-eu* (sauf : *des bleus* et *des pneus*)

- (a) Il arrive au volant d'une voiture décapotable, ses **cheveux** gris au vent.
- (b) Nos **vœux** de succès dans la poursuite de sa carrière l'accompagnent.
- (c) A l'aube, cinq ponts de la ville ont été bloqués par des autonomistes qui incendiaient des **pneus**.

1.4. Sept noms en *-ou* : *bijou, caillou, chou, genou, hibou, joujou, pou*

- (a) Yannick a très vite ressenti une irrésistible attirance pour les **bijoux** et les **cailloux** qui, bien sûr, étaient ses **joujoux** préférés.

2. Avec changement perceptible à l'oral : deux finales distinctes

2.1. Les noms en *-al* forment leur pluriel en *-aux*.

- (a) Les **chevaux** de course d'aujourd'hui appartiennent à une même famille, celle des purs-sangs.
- (b) Les **journaux** ont publié des statistiques détaillées.

5. Le pluriel des noms composés

5.1. Les noms composés qui s'écrivent en un seul mot forment leur pluriel comme les mots simples (*gendarmes*), sauf les couples *monsieur / messieurs* ; *madame / mesdames* ; *bonhomme / bonshommes* ; *gentilhomme / gentilshommes* dont le déterminant ou l'adjectif est variable.

(a) Elle se lève, hésite un instant et lance : « **Messieurs les gendarmes**, bonsoir ! »

5.2. Dans les noms composés non soudés (avec ou sans trait d'union), seuls le nom et l'adjectif peuvent prendre la marque du pluriel (*un rouge-gorge* > *des rouges-gorges* ; *un pur-sang* > *des pur-sangs*), les autres éléments demeurant invariables :

Adverbe + nom : *une arrière-boutique* > *des arrière-boutiques*

Verbe + verbe : *un laissez-passer* > *des laissez-passer*

Verbe + adverbe : *un passe-partout* > *des passe-partout*

Verbe + conjonction + verbe : *un va-et-vient* > *des va-et-vient*

5.3. Les noms composés de **deux noms** prennent chacun la marque du pluriel s'ils entretiennent un rapport d'équivalence : *une porte-fenêtre* > *des portes-fenêtres* (« des portes qui sont en même temps des fenêtres ») ; *une jupe-culotte* > *des jupes-culottes* (« pantalon très ample coupé de manière à tomber comme une jupe »).

(a) Pour des **choux-fleurs** plus blancs, ajoutez du lait à l'eau de cuisson.

Si le second nom a la fonction d'un complément introduit ou non par une préposition, il reste invariable : *un timbre-poste* > *des timbres-poste* (« des timbres de la poste ») ; *une pause-café* > *des pauses-café* (« pauses que l'on fait pour prendre le café ») ; *une eau-de-vie* > *des eaux-de-vie* (« alcool très fort à base de fruits ou de céréales »).

(b) Certains films du cinéma muet sont considérés comme des **chefs-d'œuvre** par les experts.

5.4. Les noms composés d'un **verbe** et d'un **nom**, avec trait d'union, suivent la règle des mots simples, et prennent la marque du pluriel sur le second élément¹ : *un pèse-lettre* > *des pèse-lettres* ; *un cure-dent* > *des cure-dents* ; *un abat-jour* > *des abat-jours*.

(a) Suivant les régions, les **perce-neiges** portent aussi le nom de gouttes de lait ou clochettes d'hiver.

(b) Depuis 1978, Bacchus Antique fournit des **tire-bouchons** de collection dans le monde entier.

Il en va de même des noms composés d'une **préposition** et d'un **nom** : *un après-midi* > *des après-midis* ; *un après-ski* > *des après-skis*.

(c) Cette année-là, notre pays a vécu des **avant-goûts** du changement qui ont bouleversé son mode de vie politico-économique.

¹ Remarque : cette règle obéit aux recommandations de l'Académie Française datant de 1990 (« Réforme de l'orthographe »). L'usage, cependant, tend à garder les anciennes règles, plus compliquées.

Sauf : huit noms qui forment leur pluriel en *-als* : *bals*, *calés*, *carnavals*, *cérémonials*, *chacals*, *festivals*, *réitals*, *régals*.

(c) Nous animons les **bals**, les thés dansants, les mariages, les anniversaires.

(d) Salons, expositions et **festivals** se multiplient – pour le meilleur et pour le pire.

2.2. Sept noms en *-ail* forment leur pluriel en *-aux* : *bail* > *baux* ; *corail* > *coraux* ; *émaux* ; *soupirail* > *soupiraux* ; *travail* > *travaux* ; *vantail* > *vantaux* ; *vitrail* > *vitraux*.

(a) Les **vitraux** de l'église et de la chapelle de Marguerite éblouissent par leur couleur.

(b) Elle accueille les clients, fait visiter les appartements, participe à la signature des **baux**.

Tous les autres noms en *-ail* forment régulièrement leur pluriel en *-ails*.

(c) Les **détails** de cette découverte sont disponibles sur le site Internet de la revue scientifique *Nature*.

(d) Le concours récompensera les plus beaux **épouvantails**.

3. Trois noms monosyllabiques : pluriel graphique régulier et forme orale irrégulière

Les noms *œuf* [œf], *beuf* [bœf] et *os* [ɔs] deviennent *œufs*, *beufs* et *os*, mais, au pluriel, la consonne finale n'est pas prononcée et la voyelle se ferme : *des beufs* [bø] ; *des œufs* [ø] ; *des os* [o].

4. Trois noms à double forme au pluriel, avec changement de sens

Les noms *aitel*, *ciel* et *œil* ont deux formes différentes au pluriel, qui correspondent à deux sens différents.

un aitel > *des aiteux* « des ancêtres »
> *des aiteuls* « des grands-pères » ; « des grands-parents »
> *des aiteules* « des grands-mères »

un ciel > *des cieux* « lieu surnaturel » ; « espaces infinis »
> *des ciels* « espaces visibles au-dessus de nos têtes » ; « parties d'une peinture représentant le ciel »
(et dans les mots composés, comme : *des ciels de lit*)

un œil > *des yeux* « des organes de la vue »
> *des œils* (dans les mots composés, comme : *des œils-de-beuf*)

Exercice 1 : Dans les phrases ci-dessous, soulignez les noms singuliers auxquels on ne peut pas ajouter un -s pour former le pluriel.

1. Il est environ 20 heures, le repas vient de se terminer et les membres de la famille débarrassent la table.
2. On voyait un jardin en croix, un puits au milieu, délimité en quatre parties.
3. Partout, la vie a repris son cours.
4. Il a travaillé avec les moines, creusé des rigoles et cimenté des murs, vêtu d'un bleu de travail.
5. Hier, c'était gâteau de foie de volaille avec coulis de tomate au basilic.
6. La jeune femme, qui sortait du bal des pompiers, a été admise au service de réanimation de l'hôpital de Roanne après une immersion accidentelle dans le canal.
7. Pour une raison indéterminée, le plateau s'est mis en mouvement, arrachant le portail métallique de l'entreprise, puis traversant la route pour finir sa course dans un mur en béton protégeant des installations de Gaz de France.
8. Ils sont disponibles en dix-huit coloris et portent, gravé dans l'émail ou l'inox, le nom de leur propriétaire.
9. A la sortie, il faudra penser au pourboire.
10. Elle a reçu une volée de pierres, notamment un gros caillou à la tête.

Exercice 2 : Corrigez s'il y a lieu les phrases ci-dessous.

1. Assise dans son fauteuil, un journal déplié sur ses genoux, elle sourit.
2. D'autres feux se sont déclenchés dans la journée dans le sud de la France, attisés par le vent violent.
3. Les pneus traditionnels perdent beaucoup de leurs qualités à basse température.
4. Pour les enfants, chez Disney, on trouvera des pyjamas à 13 francs, et des chemises à 4 francs.
5. Elle a refusé toute forme d'hygiène et s'est plainte de violents maux de ventre.
6. Une formule d'abonnement pour les trois réciteux est également proposée.
7. Il ne suffit pas de renvoyer un dictateur pour créer des routes, des hôpitaux et des écoles.
8. La lutte anti-spam entraîne un accroissement considérable des en-tête dans les messages.
9. Durant cette dernière décennie, les hauts-parleurs ont considérablement évolué.
10. Les ordinateurs portables sont exclus depuis peu des wagons-restaurant de Pologne.

Exercice 3 : Mettez au pluriel les syntagmes nominaux en italique et faites les accords nécessaires.

1. Le logement de Suzanne fait partie des cinq appartements du rez-de-chaussée les plus détruits.
2. *Cet incendie fut sans conséquence car il fut très vite maîtrisé grâce à certains riverains armés d'un seuil d'eau.*
3. Une torche à la main, il transmet un signal au barreur.

4. *Le résultat obtenu est au-dessus du seuil réglementaire.*

5. C'est lui qui a fait *le choix décisif* et a fait accepter *le compromis*.

6. *Le jeu* apporte aux enfants des chances supplémentaires de se développer et de s'épanouir, à condition que les parents prennent le temps de faire *le bon choix*.

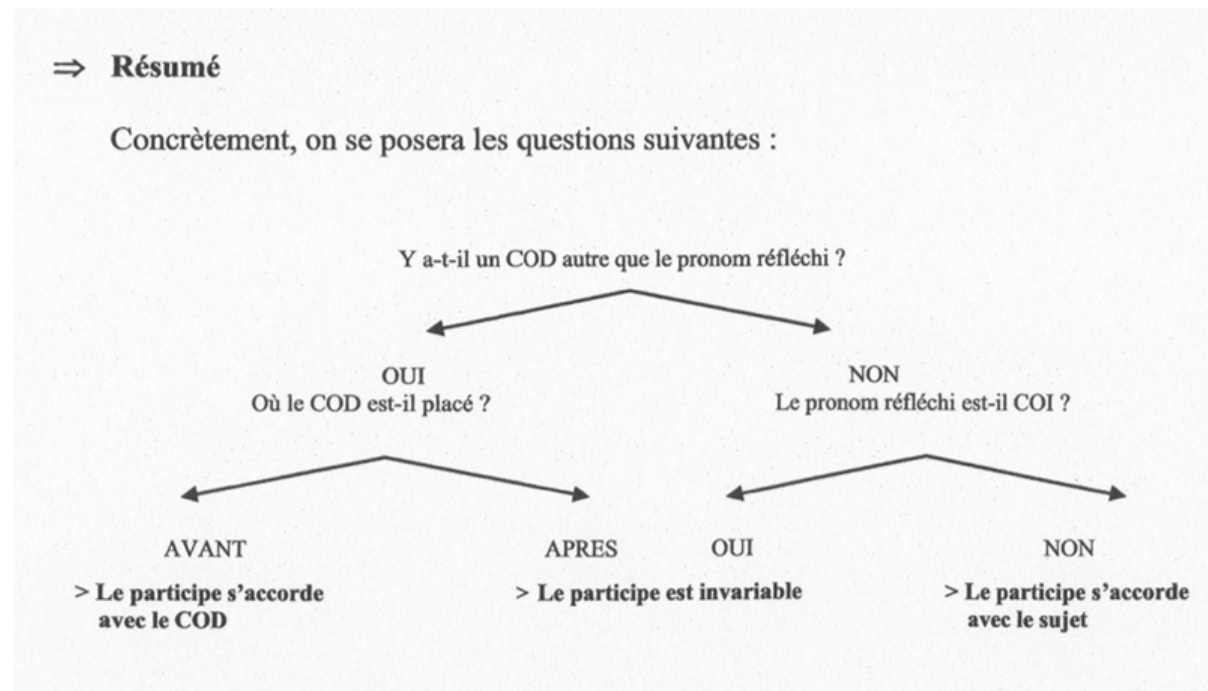
7. Max Ernst, trois cents ans plus tard, découvre *un sphinx* dans l'empreinte d'une éponge, *un oiseau* dans *le corail*, *un dragon* dans la mousse.

8. Le montage du *porte-bagage* sur le vélo lui-même peut être délicat.

9. Dans la symbolique occidentale, *l'arc-en-ciel* est souvent associé à la joie et la gaieté ou au renouvellement.

10. Un ennemi pourrait-il couler un *porte-avion* américain ?

Annexe 2b: Skupien Dekens, C., Kamber, A. & Dubois, M. (2010): Manuel d'orthographe pour le français contemporain, Neuchâtel, Université de Neuchâtel, *in* 58.



Annexe 2c: Skupien Dekens, C., Kamber, A. & Dubois, M. (2010): Manuel d'orthographe pour le français contemporain, Neuchâtel, Université de Neuchâtel, *in* 103.

Exercice 2 : Trouvez le mot mal orthographié dans les listes suivantes et corrigez-le.

1. siecle, cercle, architecte, échec _____
2. bref, chef, avec, aid _____
3. mer, cher, diver, hiver _____
4. affaire, nécessaire, fraire, salaire _____
5. jeunesse, recette, planète, dète _____

Exercice 3 : Trouvez le mot mal orthographié dans les listes suivantes et corrigez-le.

1. universel, réel, officiel, model _____
2. potentiel, anuel, industriel, actuel _____
3. elle, échelle, essentielle, hôtelle _____
4. test, pièce, ouest, gest _____
5. système, intèrne, planète, fidèle _____

Annexe 3a: Exercice à trous sur les homonymes.

EXERCICE D'ORTHOGRAPHE

1. À Strasbourg, en hiver, on peut se promener le long de l'Ill et nourrir les _____ qui, nonchalants, se laissent porter par le _____ du fleuve.
(cygnes) (cours)
2. La perspective actionnelle est une méthodologie d'enseignement des langues qui privilégie l'approche par _____. (tâches)
3. À partir du mois prochain, les enfants du quartier pourront profiter d'une nouvelle _____ de jeux. (aire)
4. Les _____ mènent une vie _____. (saints) (ascétique)
5. Dans le _____, le navire levait _____. (port) (l'ancre)
6. Nul n'est _____ ignorer la loi. (censé)
7. Venez dîner avec moi : nous ferons bonne _____. (Th 319)¹⁴ (chère)
8. Théophile est un grand consommateur de filets de _____. (sole)
9. Mon amie italienne fait de délicieuses pâtes au _____. (basilic)
10. Le jour est _____ de disparaître. (Th 322) (près de)
11. J'ai manqué mon train. La prochaine fois, je me rendrai à la gare _____.
(plus tôt)
12. Les _____ que je suis allé au théâtre me donnent envie d'y retourner. (Th 322) (quelques fois)
13. Travaillez _____ : vous réussirez mieux. (Th 322) (davantage)
14. Les deux avocats disputèrent _____ sans parvenir à un accord. (à l'envi)
15. Pour régler un tel _____, il faut aller en justice. (Th 320) (différend)
16. C'est au patron que tu _____ si tu continues à négliger ton travail.
(auras affaire)
17. _____, _____ le temps, je viendrai te _____. (ce soir)
(quel que soit), (voir)
18. C'est toi qui _____. Ne te _____ pas maintenant ! (l'as voulu)
(plaints)
19. L'olivier qu'elle _____ très vite dans _____
appartement. (s'est acheté) (croît) (son)

¹⁴

Les phrases 7, 10, 13 et 15 sont tirées de R. Thimonnier, 1974 : 319, 322 et 320, d'où les indications entre parenthèses (Th 319), (Th 322) et (Th 320).

20. Vos parents _____ sur certains points,
 _____ sur d'autres. (sont tolérants) (mais exigeants)
21. Il a prononcé un discours _____ tout le monde. (convainquant)
22. Il faut qu'elle _____ cette affaire au plus vite. (conclue)
23. Melpomène et Igor vivent en _____. (colocation)
24. L'enfant _____ sa maman, mais cette dernière n'entend pas ses
 _____. (appelle) (cris)
25. Il a assommé son agresseur d'un _____. (coup de poing)

Annexe 3b: Les performances orthographiques des étudiants allophones et endophones à l'exercice à trous portant sur les homonymes.

FRÉQUENCE DE RÉUSSITE AUX HOMONYMES					
EN FLM (30 COPIES)			EN FLES (30 COPIES)		
Mots du corpus	Graphies justes	Fréquence	Mots du corpus	Graphies justes	Fréquence
cygnes	30	1	cygnes	12	0,4
cours	24	0,8	cours	13	0,43
tâche(s)	29	0,97	tâche(s)	23	0,77
aire	29	0,97	aire	17	0,57
saints	29	0,97	saints	18	0,6
ascétique	24	0,8	ascétique	12	0,4
port	30	1	port	27	0,9
l'ancre	28	0,93	l'ancre	20	0,67
censé	19	0,63	censé	20	0,67
chère	3	0,1	chère	1	0,03
sole	21	0,70	sole	10	0,33
basilic	28	0,93	basilic	22	0,73
près	16	0,53	près	5	0,17
plus tôt	29	0,97	plus tôt	19	0,63
quelques fois	26	0,87	quelques fois	26	0,87
davantage	25	0,83	davantage	18	0,60
à l'envi	10	0,33	à l'envi	1	0,03
différend	14	0,47	différend	7	0,23
auras affaire	13	0,43	auras affaire	6	0,2
ce soir	30	1	ce soir	29	0,97
quel que soit	14	0,47	quel que soit	8	0,27
voir	29	0,97	voir	30	1
l'as voulu	12	0,4	l'as voulu	9	0,3
plains	22	0,73	plains	11	0,37
s'est acheté	25	0,83	s'est acheté	19	0,63

croît	27	0,9	croît	10	0,33
son	30	1	son	28	0,93
sont tolérants	25	0,83	sont tolérants	18	0,6
mais exigeants	24	0,8	mais exigeants	16	0,53
convainquant	17	0,57	convainquant	6	0,2
conclue	20	0,67	conclue	9	0,30
colocation	22	0,73	colocation	13	0,43
appelle	30	1	appelle	20	0,67
cris	30	1	cris	21	0,7
coup de poing	30	1	coup de poing	16	0,53
Total	814 graphies justes sur 1050		Total	540 graphies justes sur 1050	
	soit 77% de graphies justes			soit 51% de graphies justes	

Annexe 3c: Dialogue ayant servi d'amorce à l'entretien.

Au tribunal

LE PROCUREUR (*ricanant*) – Le petit Albert pourra-t-il nous expliquer pourquoi il avait écrit dans une dictée : “la tante est insupportable”? Je dis bien: la tante avec un “a”! Il ne va tout de même pas prétendre que c’est sa tente de camping qu’il trouvait insupportable ! Voyons, c’est bien de Mademoiselle Rossi qu’il s’agissait! ... Pour moi, Monsieur le Juge, l’affaire est claire: ce vaurien est un récidiviste et il mérite les travaux forcés à perpétuité!

LE JUGE – Accusé, qu’avez-vous à dire pour votre défense?

L’ACCUSÉ – Heu... C’était la faute de l’apostrophe.

LE JUGE – Quoi ?

L’AVOCAT DE LA DÉFENSE – Une simple faute d’apostrophe, Monsieur le Juge. Le texte de la dictée parlait d’une personne qui attendait depuis longtemps le retour d’un être cher et trouvait cette attente insupportable... Encore un mauvais tour de l’orthographe, hélas!

Gianni Rodari, *Histoires au téléphone*

Du lien entre la maîtrise du système phonologique et les compétences orthographiques

Saliha AMOKRANE

Université d'Alger 2

In our article, we are going to focus on the results obtained in our research which aimed at checking spelling skill of Algerian students. The corpus analysis in which we used three types of tests: dictation, writing and copying exercise allowed us to come out success as well as failure. Some mistakes made by Algerian students are common with those made by French native speakers; others are however specific to the Algerians. We have more particularly noticed the important impact that the phonology skill has on the more or less successful spelling practices in all suggested tests.

L'orthographe française qui est considérée comme un système "opaque" ou "profond" dans lequel on représente, outre le niveau phonologique, les aspects morphologiques et syntaxiques, pose de sérieux problèmes aux scripteurs francophones natifs et son apprentissage est un parcours semé d'embûches comme en témoignent ces propos d'une jeune lycéenne française (élève du lycée professionnel en 2^{ème} année de CAP)¹ rapportés par A. Millet, V. Lucci & J. Billiez (1990: 230): "moi je connais pas les règles d'orthographe, alors je sais pas j'en connais certaines, mais je les connais pas toutes. Moi je trouve que c'est nul moi ce truc, c'est nul! Pourquoi pas écrire les mots, les choses, comme on les entend, c'est tout hein". Si le problème est posé pour des francophones natifs, il ne peut que l'être de manière encore plus ardue pour des étudiants dont la langue maternelle n'est pas le français, comme c'est le cas des étudiants algériens qui ont comme langue maternelle le berbère et/ou l'arabe dialectal et comme première langue d'enseignement l'arabe dit classique. C'est ce qui a motivé notre travail de recherche en doctorat, qui visait à vérifier la maîtrise de l'orthographe par des étudiants algériens du département de français de deux universités algériennes – Alger et Constantine – et à déterminer les sources de leurs problèmes. Dans le cadre de cet article, nous partons des résultats de cette recherche.

Le choix de ce public se justifiait par le fait qu'il s'agissait d'étudiants (des quatre années de licence)² c'est-à-dire d'apprenants qui en principe

¹ CAP: Certificat d'aptitude professionnelle.

² A l'époque où nous avons réalisé notre recherche, le système LMD (Licence – Master – Doctorat) n'avait pas encore été introduit dans les départements de français des deux universités.

n'étaient plus en phase d'apprentissage de l'orthographe. En effet, si l'on se réfère aux programmes d'enseignement du français en Algérie, on se rend compte que l'enseignement de l'orthographe n'est quasiment plus pris en charge à partir de la 8^{ème} année fondamentale³ (soit la 5^{ème} année d'enseignement du français)⁴. De fait, certaines unités didactiques comportant un volume horaire global de 12 heures ne prévoient aucun contenu pour la séance d'orthographe à laquelle on consacre normalement 1 heure. Au lycée, l'orthographe n'apparaît plus comme activité dans le programme, même si les concepteurs de ce dernier précisent que "la liste des activités et notions figurant dans le tableau des contenus de l'UD⁵ n'est pas exhaustive"⁶. La deuxième raison de notre choix est liée au fait que ces étudiants préparent une licence de français et donc qu'un grand nombre d'entre eux seront appelés eux-mêmes à enseigner la langue française et par la même occasion l'orthographe.

1. Le corpus

Sachant que, comme le souligne Brissaud (1998: 67), "la performance orthographique varie d'une tâche à l'autre et qu'il existe, par exemple, un décalage entre la performance orthographique en dictée et celle en rédaction", nous avons proposé trois types d'épreuves pour recueillir notre corpus: une dictée, un exercice de copie et une rédaction.

1.1 La dictée

Cette première épreuve concerne trois des ensembles dégagés par Catach, Duprez & Legris (1980: 13-15) dans leur typologie des erreurs, à savoir les erreurs à dominante phonétique, les erreurs à dominante phonogrammique et des erreurs portant sur les homophones.

³ L'école fondamentale a été mise en place par l'ordonnance du 16 avril 1976 et comprend les neuf premières années du système éducatif algérien. Les neuf années se répartissent en 3 cycles de 3 ans:

Le 1^{er} cycle comprend les 1^{ère}, 2^{ème} et 3^{ème} années fondamentales.

Le 2^{ème} cycle comprend les 4^{ème}, 5^{ème} et 6^{ème} années fondamentales.

Le 3^{ème} cycle comprend les 7^{ème}, 8^{ème} et 9^{ème} années fondamentales.

⁴ L'enseignement du français commençait en 4^{ème} année fondamentale. Dans le rapport final de la commission nationale de réforme du système éducatif installée en mai 2000, il est préconisé de réintroduire le français dès la 2^{ème} année du cycle primaire. Mais, à l'heure actuelle il commence en 3^{ème} année sous prétexte que "l'élève doit d'abord se familiariser avec sa langue nationale pendant les deux premières années de sa scolarisation".

⁵ L'unité didactique.

⁶ Programme de français de l'enseignement secondaire, Office national des publications scolaires, 1991-1992.

1.1.1 La zone des phonogrammes

Sachant que selon les études menées sur des apprenants ayant le français comme langue maternelle la phonographie se met en place en premier, il nous a semblé que la maîtrise de celle-ci par des apprenants n'ayant pas le français comme langue maternelle méritait d'être vérifiée, d'autant plus que la zone des phonogrammes représente la partie centrale du système graphique.

Notre hypothèse était que, du fait de la nature des systèmes linguistiques en présence, cette zone du plurisystème graphique risquait de poser problème et cela pour différentes raisons.

Tout d'abord, puisque la maîtrise des écritures alphabétiques exige la reconnaissance des phonèmes, nous avons pensé que des difficultés au niveau phonique pouvaient avoir des répercussions sur le maniement du système graphique. Or, les systèmes vocaliques arabe et français sont particulièrement différents. En effet, face à la richesse du système vocalique français, qui constitue une source majeure de problèmes dans la mesure où, comme l'affirme Argod-Dutard (1996: 58) "il est difficile sur les plans auditif et phonatoire de distinguer quatre degrés surtout à l'arrière de la bouche, de bien maintenir deux séries vocaliques d'avant", nous avons, en arabe (qui est la première langue apprise par nos étudiants) un système vocalique qui ne compte que six voyelles dont trois brèves /a/-/i/-/u/ et les trois longues correspondantes /a:/-/i:/-/u:/⁷. De ce fait et dans la mesure où, comme l'a montré Troubetzkoy, "le crible phonologique" de la langue maternelle a tendance à réduire considérablement les capacités de discrimination et de catégorisation phonémique des non natifs, il est fort probable que les apprenants butent sur les phonèmes vocaliques et donc qu'un grand nombre d'erreurs soit d'origine phonétique.

Ensuite, le français, du fait de la pluralité des transcriptions potentielles d'un même phonème ainsi que de la pluralité des valeurs que peut avoir une même graphie selon le contexte d'apparition, est considéré comme non biunivoque, ce qui implique qu'il est possible de produire dans cette langue des formes phonologiquement plausibles mais qui, au regard de la norme orthographique, seront considérées comme inadmissibles (ex: "otomobile"). A l'inverse, nous avons en arabe, une correspondance entre phonie et graphie presque générale. En effet, selon Koughougli, cité par Ghellaï (1997: 50), "le principe fondamental de l'écriture arabe est que la graphie normale d'un mot reflète exactement sa prononciation". En fait, le système graphique de l'arabe fait coexister deux formes d'écriture: l'une dite "pointée" qui permet de noter aussi bien les voyelles que les

⁷ Koughougli, cité par Ghellaï (1997: 49) précise que "le petit nombre de voyelles de l'arabe fait que chacune d'entre elles dispose d'un large espace, ce qui permet des réalisations phonétiques variées".

consonnes et qui est donc très proche de la prononciation et l'autre "non pointée" qui ne note pas les voyelles brèves et qui note donc moins que la prononciation. De ce fait, nous avons supposé que l'existence, en français, de plusieurs graphèmes pour transcrire un seul et même phonème, voire de graphèmes qui ne notent aucun phonème (le français note plus que ce qui est prononcé) va être une source de difficultés pour des arabophones.

1.1.2 La zone des logogrammes

Cette zone du plurisystème graphique n'est pas centrale. Toutefois, il nous a semblé bon de la contrôler dans la mesure où nous avons fait l'hypothèse que la particularité du système graphique arabe que nous venons de mentionner, à savoir la correspondance quasi parfaite entre phonie et graphie, allait entraîner des difficultés à ce niveau également. En effet, nous avons supposé que cette zone qui représente, même pour des francophones natifs, une source majeure de problèmes allait constituer pour les étudiants algériens une pierre d'achoppement supplémentaire.

Pour vérifier ces différentes hypothèses, la dictée proposée comportait plusieurs parties:

La première partie composée d'une dictée de termes isolés visait à vérifier la maîtrise de certaines oppositions phonologiques que nous avons considéré, au regard des systèmes vocaliques en présence, comme pouvant être à l'origine de nombreuses erreurs à dominante phonétique. Par exemple, pour les voyelles orales les oppositions /i/~E/ (ex: *précis*); /i/~y/ (ex: *illusion*); /u/~o/ (ex: *fourreau*)⁸; /ø/~o/ (ex: *eurovision*); /œ/~ɔ/ (ex: *horreur*)⁹; et pour les voyelles nasales, inexistantes en arabe, les oppositions /ã~/ĩ/ (ex: *attention*); /ẽ~/ĩ/ (ex: *tympan*)¹⁰. Nous avons dicté des séries de 5 termes pour chaque opposition¹¹. Dans chaque terme proposé apparaissent les deux unités susceptibles d'être confondues, l'objectif étant de savoir si les étudiants parviennent à les distinguer et à les opposer fonctionnellement.

Dans la deuxième partie de la dictée de termes isolés, nous voulions contrôler la maîtrise des différents allographes (même si certaines des

⁸ Dans ces trois séries d'oppositions, il s'agissait de contrôler la maîtrise de l'opposition entre un phonème existant en arabe et un phonème inexistant.

⁹ Dans ces deux séries, les deux phonèmes en opposition sont inexistantes en tant que phonèmes.

¹⁰ Tous ces exemples sont extraits de notre corpus.

¹¹ Après coup, il nous a semblé que pour le contrôle des oppositions phonologiques, il aurait été préférable de partir de non-mots car, dans ce cas, le recours nécessaire à la voie phonographémique pour transcrire les unités aurait permis aux confusions de se manifester de manière plus systématique. En effet, nous avons remarqué que pour la même opposition le nombre d'erreurs varie selon la fréquence du mot proposé et la familiarité des étudiants avec ce dernier.

formes proposées, telles que "paon" pourraient être considérées comme logogrammiques) correspondant à 14 phonèmes (soit 9 consonnes: /s/, /ž/, /k/, /g/, /z/, /j/, /t/, /f/ et /R/ et 5 voyelles: /ẽ/, /O/, /ã/ /E/, /i/). A ce niveau, nous avons proposé autant de termes que de graphèmes correspondant à chacun des phonèmes. Pour le choix des différents allographes, nous nous sommes référée aux grilles analytiques proposées par Gey (1987: 25-33).

Dans la troisième partie de la dictée, qui était une dictée de phrases, nous nous proposons de contrôler la maîtrise de certains logogrammes, grammaticaux pour la plupart, à savoir: se/ce; ces/ses/c'est/s'est; et/est; a/à; on/ont; son/sont; sans/sang; tous/tout; ou/où; si/s'y; la/là.

1.2 *La rédaction*

Cette deuxième épreuve était une question ouverte dans laquelle les étudiants avaient toute liberté de s'exprimer à leur guise mais en même temps toute latitude de n'utiliser que les unités ne comportant pas de difficultés orthographiques particulières (au niveau phonogrammique par exemple) et les structures ne nécessitant pas de calculs trop complexes (au niveau morphogrammique). Les étudiants risquaient donc d'éviter, autant que possible, les formes qui leur posaient problème. Une des stratégies d'évitement fréquemment utilisée a été de limiter la réponse à une phrase, de manière à prendre le minimum de risques et donc de s'exposer le moins possible (à ce propos, nous avons d'ailleurs noté que 23,20% des étudiants ont produit moins de 5 lignes)¹².

De ce fait, les données que nous pouvions rassembler à travers cette épreuve seule risquaient d'être insuffisantes car, si elle nous permettait de repérer un certain nombre d'erreurs, elle risquait également de laisser se manifester des "fautes invisibles" ou "fautes en négatif"¹³, selon la terminologie de Rojas (1971: 62).

Toutefois, la rédaction, dans la mesure où elle était combinée à deux autres épreuves très différentes (la dictée et la copie), nous semblait intéressante car elle allait nous permettre de mieux faire la part entre erreurs de compétence (qui se retrouvent quelle que soit la nature de l'épreuve) et erreurs de performance (liées à la tâche).

¹² Le nombre total d'étudiants examinés est de 292 dont 156 à Alger et 136 à Constantine. Ils sont répartis comme suit: 1^{ère} année: 45 à Alger et 37 à Constantine; 2^{ème} année: 38 à Alger et 36 à Constantine; 3^{ème} année: 24 à Alger et 32 à Constantine; 4^{ème} année: 49 à Alger et 31 à Constantine.

¹³ On parlera de "fautes en négatif" lorsqu'un apprenant qui ignore totalement ("usage nul") ou partiellement une forme ne l'utilise pas soit par méconnaissance soit par autocensure (peur de l'erreur) ou stratégie d'évitement.

Pour recueillir les productions des étudiants, nous avons posé la question suivante: "Vous avez été orientés en licence de français. Etes-vous satisfaits de cette orientation? Justifiez votre réponse".

Cette question présentait l'intérêt de concerner tous les étudiants et donc de permettre à chacun d'entre eux d'avoir quelque chose à dire. En effet, sachant que l'orientation est effectuée par ordinateur sur la base de 15 propositions faites par le nouveau bachelier mais en fonction des places disponibles dans chaque filière, il arrive rarement que le bachelier soit orienté vers les filières qui représentent ses premiers choix (ceci n'est pas vrai pour les meilleurs d'entre eux puisque les élèves ayant obtenu une mention "bien" ou "très bien" au bac ont la possibilité de ne faire qu'une seule proposition qui est automatiquement acceptée). L'orientation représente donc un véritable problème qui touche la plupart des étudiants.

Le choix de cette thématique, directement en rapport avec la vie de ces étudiants, se justifiait encore par le fait que, comme cela a été mis en évidence par nombre d'études, lorsque le thème de la production est trop étranger au producteur du texte, il peut constituer un facteur paralysant. De fait, selon Fayol (1984: 66), "l'écrivain "jongle" avec les contraintes inhérentes au "fond" et à "la forme"" ou pour reprendre une terminologie plus récente, passer d'opérations de "bas niveau" à des opérations de "haut niveau". Ceci conduit le scripteur à "négliger", bien souvent, l'orthographe ou plutôt à laisser se manifester "un dysfonctionnement dû à l'incapacité de contrôler aussi"¹⁴ le code orthographique quand le sujet à traiter est trop paralysant" (Dabin & Baudry, 1985: 38).

Par le thème choisi, cette épreuve représentait donc une situation relativement "favorable" pour l'étudiant. Mais, en fait, dans la mesure où il devait donner une opinion personnelle et la justifier, son attention ne pouvait pas être mobilisée exclusivement par l'orthographe, comme cela pouvait être le cas dans le premier test, à savoir la dictée.

1.3 *La copie*

La troisième épreuve était la copie d'un texte. Les erreurs de copie sont souvent jugées impardonnables car recopier un texte sans contrainte de temps semble être un exercice très facile. Or, ces erreurs existent et nous pensons qu'elles peuvent, également, être liées à des difficultés particulières que rencontrent les élèves. Dans ce cas, leur étude permet de mieux savoir quels sont les problèmes posés par l'apprentissage de l'orthographe et de repérer les zones du plurisystème qui ne sont pas maîtrisées. De fait, nous avons fait l'hypothèse que certaines zones sont si peu maîtrisées qu'elles ne sont même pas perçues par les apprenants (les

¹⁴ C'est nous qui soulignons.

accents, la ponctuation, les majuscules en particulier). Nous avons donc essayé à travers ce test de vérifier le concept de "cécité sélective"¹⁵ qui nous a semblé pouvoir caractériser cette attitude des étudiants face à certaines caractéristiques du texte à recopier.

2. Les résultats

L'analyse tant quantitative que qualitative des productions des étudiants obtenues à travers les trois tests nous a permis de noter qu'à tous les niveaux examinés (phonétique, phonogrammique, logogrammique) erreurs et réussites coexistent.

2.1 Les réussites

Nous avons pu noter que les étudiants sont conscients du caractère non biunivoque du système graphique français. En effet, à l'exception de "gg", toutes les graphies susceptibles de transcrire les différents phonèmes et que nous avons contrôlées sont réalisées, qu'ils s'agissent des plus fréquentes comme les archigraphèmes (ex: IN pour transcrire /ɛ̃/, O pour transcrire /o/) ou des plus rares comme les sous-graphèmes (ex: "ain" ou "im" pour transcrire /ɛ̃/) voire les graphies logogrammiques (ex: "aon" pour transcrire /ã/).

Par ailleurs, nous avons pu constater que ce sont le plus souvent les archigraphèmes qui sont utilisés pour transcrire les différents phonèmes que nous avons contrôlés, ce qui signifie que les étudiants sont conscients de la fréquence relative des différents graphèmes. Nous n'avons relevé qu'un seul cas où ceci ne se vérifie pas, celui de l'archigraphème AN qui a souvent été remplacé par "en". Or, nous savons qu'en fait ces deux graphèmes ont une fréquence quasi identique en français et que si Catach a retenu AN comme archigraphème c'est parce que c'est la graphie la moins ambiguë ("en" pouvant transcrire également /ɛ̃/ ex: "viens" ou /ɛn/ ex: "spécimen").

Enfin, toujours en ce qui concerne les phonogrammes, nous avons pu constater que les scripteurs utilisent des unités définies avec des règles phonogrammiques pratiquement toujours motivées judicieusement par rapport au système graphique français. Les cas de graphies hors système que nous avons pu relever telles que "ku" et "cku" pour transcrire /k/ ou "quu" et "qu" pour transcrire /ky/, sont extrêmement rares.

Pour les consonnes doubles, même s'il y a encore de très nombreuses erreurs d'adjonction et de suppression, nous avons néanmoins pu noter que les étudiants ont mémorisé un certain nombre de règles quant à la

¹⁵ Ce concept que nous proposons est élaboré sur le modèle du concept de "surdité sélective" utilisé en phonétique.

nature des consonnes qui peuvent être doublées et quant à la place des consonnes doubles. En effet, d'une part les étudiants n'ont jamais doublé les consonnes qui ne le sont pas en français et d'autre part, nous n'avons relevé aucun cas de consonnes doubles à l'initiale et relativement peu en finale absolue (ex: *Ecureuill*, *giraff*) et après une autre consonne (ex: *sangllie*).

Pour les consonnes finales muettes, nous pouvons dire qu'à défaut de les réaliser toujours avec pertinence, les étudiants en manifestent le souci de façon récurrente. En effet, alors qu'elles n'existent pas en arabe, il arrive fréquemment qu'elles soient notées même dans des contextes où elles ne sont pas requises. De plus, nous constatons que les étudiants n'utilisent comme consonnes finales muettes que des consonnes qui jouent fréquemment ce rôle en français, telles "t" et "d", essentiellement dans des formes comme *pent* pour "paon", *pigeant* pour "pigeon", *jauard* pour "jaguar".

Par ailleurs, comme pour les consonnes finales muettes, les étudiants manifestent, également, un réel souci pour le "e" final qui a beaucoup plus souvent été ajouté (après le "r", le "l", le "z", le "f" mais également après le "i") que supprimé.

En ce qui concerne les logogrammes, enfin, les scripteurs utilisent généralement les graphies conventionnelles pour les différents homophones, même s'ils ont tendance à sur-utiliser une graphie dans chaque couple lorsque ceux-ci fonctionnent en couple. Exemple: pour les homophones en /u/ nous avons une domination de la forme "ou", de la même manière pour les homophones en /sə/, nous avons domination de la forme "se".

2.2 Les erreurs

A côté des réussites dont nous venons de faire état, nous avons, également, pu observer un certain nombre d'erreurs. Certaines ne sont pas spécifiques aux étudiants algériens, d'autres, par contre, le sont.

2.2.1 Les erreurs non spécifiques aux étudiants algériens

Tout d'abord, nous avons noté un grand nombre d'erreurs sur les diacritiques qui sont le plus souvent supprimés. Ce type d'erreur a déjà été mis au jour par Lucci & Millet (1994) chez les scripteurs francophones natifs.

Nous avons également relevé de nombreux cas de confusion entre les différents allographes susceptibles de transcrire un même phonème. Ceci a été, également, signalé en FLM par Lucci & Millet (1994: 101). En effet, ils ont mis en évidence des hésitations sur la transcription du /s/, du /E/, du

/ã/ et plus sporadiquement du /i/ et du /o/. Ils précisent que "ces zones présentent sans doute des possibilités concurrentielles de transcription d'un son trop nombreuses pour que tous les scripteurs, dans des situations où les contraintes sont fortes, puissent réaliser des graphies normées pour toutes les occurrences de ces sons".

Mais, en ce qui concerne les étudiants algériens, l'éventail des erreurs est plus large. En effet, nous remarquons que les étudiants algériens ont des difficultés à transcrire non seulement les phonèmes cités par Lucci & Millet mais également dans l'ordre croissant de difficulté, le /t/, le /f/, le /ʒ/, le /z/, le /ẽ/, le /g/, le /j/, le /k/ et le /R/.

Par ailleurs, nous avons noté des erreurs liées à la non maîtrise des valeurs de position de certains graphèmes tels que le "s" intervocalique, le "c" devant "a", "o", "u", le "g" devant "a", "o", "u", le "ç" devant "i" qui se retrouvent également chez les jeunes scripteurs francophones natifs.

Enfin, nous avons constaté des erreurs sur les homophones qui sont sensiblement les mêmes que celles qui ont été dégagées par des chercheurs travaillant sur des scripteurs francophones natifs. En effet, comme nous l'avons nous-même constaté, ils ont noté une tendance à privilégier une graphie (souvent la plus courte) dans chaque couple.

2.2.2 Les erreurs spécifiques aux étudiants algériens

Nous avons pu relever l'impact très important de la maîtrise du système phonologique sur les réussites plus ou moins grandes en orthographe.

Tout d'abord, nous avons constaté que les erreurs à dominante phonétique sont omniprésentes puisqu'elles se retrouvent aussi bien en dictée, en copie qu'en rédaction. Ceci nous permet donc d'affirmer qu'il s'agit d'erreurs de compétence et non pas d'erreurs de performance.

Pour la dictée, nous avons remarqué qu'il y avait systématiquement confusion entre les différentes unités que nous avons évaluées. C'est ainsi que nous avons obtenu les réalisations suivantes: "pricis" ou "pressée" pour précis; "attention" ou "attention" pour attention; "boucoup" ou "beaucaup" pour beaucoup; "eupheri" ou "auphorie" pour euphorie; "heureure" ou "aurore" pour horreur; "initile" ou "unutil" pour inutile; "pintain" ou "panten" pour pantin. Nous avons constaté également que la confusion se faisait non seulement lorsque les unités apparaissaient dans le même mot (ce qui représente un degré supérieur de difficulté) mais également lorsqu'une seule des deux unités apparaissait (ex. "privantion" pour prévention; "euforé" pour euphorie; "ortiquilteur" pour horticulteur; "infonteré" pour infanterie).

Parmi toutes les oppositions évaluées, ce sont les oppositions /ã/~/õ/, /i/~/E/ et /i/~/y/ qui posent le plus de problèmes.

Nous retrouvons des erreurs de type phonique dans l'exercice de copie également. Celles-ci se manifestent sous forme de confusions, de suppressions, d'adjonctions ou de métathèses. Les confusions, qui sont les erreurs les plus fréquentes, portent essentiellement sur deux oppositions qui ont déjà été identifiées comme source majeure de problèmes en dictée à savoir l'opposition /ã~/ĩ/ (à travers les formes "compagne" pour campagne, "habitonts" pour habitants, "tromblent" pour tremblent...) et /o~/u/ (dans les formes "volumineuse" pour volumineuse, "autobus" pour autobus, "coullines" pour collines et inversement "s'accropir" pour s'accroupir...). Par contre, les oppositions /i~/E/ et /i~/y/, qui font partie des oppositions les moins bien maîtrisées en dictée, n'apparaissent pas comme source importante de problème en copie (le nombre d'erreurs relevées restant faible).

En rédaction, nous avons également relevé des erreurs de type phonique qui se sont manifestées sous forme de confusions, de suppressions, de nasalisations/dénasalisation, d'adjonctions, de métathèses. Toutefois, ce sont essentiellement les confusions qui dominent et, majoritairement, celles-ci concernent les oppositions qui ont été relevées comme posant problème en dictée et en copie à savoir dans l'ordre décroissant les trois paires /i~/E/, /ã~/ĩ/ et /o~/u/. Les formes fautives relevées sont pour /i~/E/: "baccalauriat" pour baccalauréat, "amiliore" pour améliorer ou inversement "enrêchir" pour enrichir, "letterature" pour littérature...; pour /ã~/ĩ/: "volanté" pour volonté, "repandre" pour répondre. Et inversement "fronçaise" pour française, "longue" pour langue...; pour /o~/u/ "volu" pour voulu, "govermentale" pour gouvernementale et inversement "époucrite" pour hypocrite, "voulenté" pour volonté...

La forme d'erreur la plus fréquente, quel que soit le type d'épreuve, semble bien être la confusion; mais ce qui mérite d'être signalé, c'est que celle-ci ne se fait pas nécessairement au détriment du phonème absent en arabe. C'est ainsi, par exemple, que pour l'opposition /i~/E/ la confusion entre les deux unités aboutit au remplacement du /i/ (qui existe en arabe) par le /E/ (absent en arabe en tant que phonème).

Les erreurs à dominante phonétique se manifestent également, même si c'est de façon plus sporadique, par la confusion entre certaines consonnes, en particulier /s~/z/ et /p~/b/.

Par ailleurs, nous avons remarqué que la non maîtrise des oppositions phonologiques perturbe aussi bien le fonctionnement des phonogrammes que des logogrammes. De fait, dans les deux cas le champ des possibles est élargi. C'est ainsi que pour transcrire un phonème les étudiants doivent choisir non seulement entre les différents allographes mais également entre ces allographes et d'autres graphèmes du fait de la confusion phonétique (ex: pour transcrire le /E/, le "i" va entrer en concurrence avec

tous les allographes susceptibles de transcrire ce phonème); de la même manière, pour les logogrammes la confusion se fera non seulement entre les formes homophones mais également entre les formes qui fonctionnent comme telles du fait de la confusion au niveau phonique (ex: le "ont" sera aussi souvent confondu avec "on" qu'avec "en", et bien plus, le "on" est plus souvent confondu avec "en" qu'avec "ont").

Les difficultés au niveau phonique se manifestent également à travers une tendance lourde à modifier la structure syllabique des items dictés. C'est ainsi que nous avons relevé des cas de suppression des groupes consonantiques consonne+liquide qui sont absents en arabe (ex: "embillant" pour embryon; "hordeuve" pour hors-d'œuvre...); des cas de suppression ou de modification des syllabes composées d'une voyelle seule, cas de figure absent également en arabe où toute voyelle est portée par une consonne (exemple de suppression: "cureuil" pour écureuil; exemple de modification par adjonction d'une consonne initiale: "posé" pour oser...).

3. Propositions

Partant du constat que les difficultés au niveau phonique sont omniprésentes et qu'elles sont source de problèmes en orthographe, nous avons pu mettre en évidence la nécessité, pour accroître et développer les compétences orthographiques d'apprenants arabophones, de renforcer la maîtrise du système phonologique plus particulièrement du système vocalique puisque, malgré la présence d'erreurs concernant les consonnes, ce sont essentiellement les voyelles qui ont posé problème. Nous avons proposé de prolonger l'enseignement de la phonétique car il nous semble tout à fait utopique de considérer que le système vocalique français dans toute sa complexité puisse être maîtrisé par des arabophones après deux années seulement d'enseignement. En effet, dans les programmes de 5^{ème} AF (2^{ème} année d'enseignement du français) il est bien question de poursuivre l'acquisition du système phonétique, d'éduquer l'écoute attentive d'énoncés de plus en plus longs, de poursuivre l'entraînement à l'effort articulatoire mais dans le guide du maître il est spécifié qu' "on reprend (pas systématiquement) les oppositions qui ne sont pas encore maîtrisées du point de vue de la prononciation". Or, si on ne les reprend pas, il est difficile d'imaginer que les problèmes peuvent alors se résoudre. Et en 6^{ème} AF, par contre, il n'est plus question de phonétique.

Cet enseignement doit, dans un premier temps, viser le renforcement des capacités de discrimination auditive (sans lesquelles il semble difficile de parvenir à une précision articulatoire). Selon le niveau envisagé, les activités devraient viser le repérage des phonèmes isolés ou celui des phonèmes en opposition. Cependant, étant donné les spécificités du

système vocalique arabe (chacune des voyelles est susceptible de recevoir diverses réalisations phonétiques selon le contexte d'apparition: par exemple le /i/ sera réalisé [i] dans un contexte non emphatique et [e] dans un contexte emphatique, de la même manière le /u/ sera réalisé [u] dans un contexte non emphatique et [o] dans un contexte emphatique)¹⁶, il nous semble essentiel d'insister sur les exercices de commutation entre les phonèmes le plus souvent confondus pour faire prendre conscience de l'impact que peut avoir, sur le sens du message, le transfert en français de cette non discrimination phonologique du système arabe.

Ensuite, il faudrait viser l'amélioration de l'articulation en faisant travailler les oppositions les plus difficiles à maîtriser eu égard aux systèmes phonologiques en présence, à travers des exercices de phonétique corrective. Bien évidemment cela nécessite de mettre à la disposition des enseignants les moyens nécessaires car il paraît malaisé d'envisager ce type d'objectif avec des classes surchargées (40 élèves minimum) et souvent de niveau très hétérogène.

Toutefois, sachant que les exercices qui ciblent une difficulté peuvent être réussis alors même que les problèmes subsistent et qu'ils peuvent resurgir dès lors qu'il faut penser au message et à sa réalisation linguistiquement correcte tant du point de vue syntaxique que lexical et phonétique, il nous semble impératif de multiplier les occasions de production spontanée pour permettre aux apprenants de mettre en pratique leurs connaissances phonétiques. Il y a donc lieu de renforcer le volet oral de l'enseignement de la langue, insuffisamment pris en charge dans l'enseignement en Algérie à tous les niveaux d'enseignement, du primaire à l'université, et cela aussi bien en réception qu'en production.

Conclusion

En conclusion, nous reprendrons les termes de Mahmoudian cité par Catach, Duprez & Legris (1980: 5), selon qui "les variations d'orthographe qui sont la contrepartie d'une différence phonique sont mieux maîtrisées car plus directement accessibles à l'intuition du sujet" et nous ajouterons que cela implique que les différences phoniques en question soient bel et bien perçues ou, en d'autres termes, que le système phonologique de la langue cible soit dominé car si tel n'est pas le cas, il paraît difficile de viser une bonne compétence orthographique. Pour que le système phonologique soit maîtrisé, il y a lieu – et cela est selon nous tout à fait impératif – de redonner à la phonétique une place de choix dans l'enseignement des langues étrangères. Cependant, l'objectif minimal (une intelligibilité confortable) souvent visé en phonétique nous semble insuffisant. En effet, si la confusion entre /i/ et /E/ peut ne pas gêner la communication à l'oral

¹⁶ /sil/ "coule", /ʔefl/ "garçon"; /kul/ "mange", /ʃof/ "laine".

puisque le contexte dans lequel est produit l'énoncé peut lever l'ambiguïté, par contre elle sera source de difficulté orthographique à l'écrit.

Bibliographie

- Amokrane, S. (2006): L'orthographe française: sa pratique et son enseignement en Algérie. Thèse de doctorat (Université d'Alger).
- Amokrane, S. (2007): Pour améliorer la maîtrise de l'orthographe en FLE/FLS... renforcer l'enseignement de la phonétique. In: les Annales de l'université d'Alger, n°17 tome, 31-37.
- Amokrane, S. (2009): Apprentissage de l'écriture en contexte plurilingue: Problèmes liés à l'acquisition de l'orthographe. In: Synergies Algérie n°6, 71-77.
- Argod-Dutard, F. (1996): Eléments de phonétique appliquée. Paris (Armand Colin).
- Brissaud, C. (1998): L'imparfait et ses concurrents du CM2 à la troisième. In: Le français aujourd'hui n°122, 62-70.
- Catach, N. (1998): L'orthographe, Que sais-je?, 8^{ème} édition. Paris (PUF).
- Catach, N., Duprez, D. & Legris, M. (1980): L'enseignement de l'orthographe. Paris (Nathan).
- Champagne-Muzar, C.S. & Bourdages, J. (1998): Le point sur la phonétique. Paris (CLE International).
- Dabin, D. & Baudry, J. (1985): Orthographe de lycéens. In: Pratiques n°46, 35-40.
- Fayol, M., (1984): Pour une didactique de la rédaction. Faire progresser le savoir psychologique et la pratique pédagogique. In: Repères n°63, INRP, 65-69.
- Gey, M. (1987): Didactique de l'orthographe française. Paris (Nathan).
- Ghellaï, M. (1997): Analyse des erreurs et des représentations orthographiques du français par des intellectuels arabophones. Thèse de doctorat (Université de Grenoble).
- Gruaz, C. (1985): Phonèmes, graphèmes, morphèmes. In: Pratiques n°46, 97-107.
- Hermeline, L. (2001): Enseigner la phonétique: oui mais comment? In: Le français dans le monde n°318, 27-28.
- Honvault, R. (1995): Statut linguistique et gestion de la variation graphique. In: Langue française n°108. Paris (Larousse), 10-17.
- Lucci, V. & Millet, A. (1994): L'orthographe de tous les jours: enquête sur les pratiques orthographiques des Français. Paris (H. Champion, Coll. "Politique linguistique 4").
- Millet, A., Lucci, V. & Billiez, J. (1990): Orthographe mon amour! PUG.
- Rojas, C. (1971): L'analyse des fautes. In: Le français dans le monde n°81, 58-63.

De la phonologie à la graphie vers leur synergie: une expérience tchèque

Danièle GEFFROY KONŠTACKÝ & Sylva NOVÁKOVÁ

Section de français, Faculté de Pédagogie, Université Hradec Králové,
& Fakulta pedagogická ZČU v Plzni, KRF

In Czech the written form of the language is mainly phonological and learners find it difficult at first to handle with the French system. The authors noticed that the methodological approach of these last twenty years did not help improving that aspect of the foreign language acquisition in secondary schools. On the basis of a short text dictated to their first year students they study the reasons why, besides having better communicative competences, the students still have problems in spelling. The authors are convinced that the link between the phonological and graphical systems of the language should be taught from the beginning of the studies to avoid an appropriation of inaccurate sequences of words and to improve the understanding and writing of meaningful sequences. They recommend an audio-oral introduction to the new language, where the differences between the two phonological systems are taken into account, and suggest the use of manuals that are conceived in that sense.

Introduction

Savoir appliquer une orthographe correcte est sans aucun doute l'une des compétences qui se révèlent les plus difficiles à acquérir, que ce soit en langue maternelle ou en langue étrangère. Les technologies de l'information et de la communication n'ont pas encore apporté de solution convaincante dans ce domaine. Or, l'orthographe fait partie des compétences "stratégiques" dans le contexte de la communication écrite (Kross, 2011), le sujet mérite donc d'être traité avec une attention particulière.

Le présent article se veut une contribution au débat autour de la méthodologie qui serait susceptible d'aider les apprenants à comprendre le système de fonctionnement de la langue française et leur permettrait de développer des habitudes d'apprentissage efficaces.

Les réflexions des auteures de cet article sont le reflet de leurs observations auprès de groupes où les apprenants sont constamment exposés aux différences multiples de deux systèmes linguistiques, le système français et le tchèque.

1. Caractéristiques principales du tchèque et du français

Pour mieux comprendre les interférences potentielles qui peuvent surgir pendant l'apprentissage du français par les apprenants tchèques, il nous paraît utile de présenter les principales caractéristiques des deux langues ainsi que leurs bases phonético/phonologiques.

	tchèque	français
classification généalogique	langue indo-européenne, groupe slave occidental	langue indo-européenne, groupe roman occidental
type morphologique	plus synthétique que le français	plus analytique que le tchèque
morphologie nominale	plus riche qu'en français, 7 cas grammaticaux	moins riche qu'en tchèque, absence de cas grammaticaux
morphologie verbale	moins riche	plus riche

Tableau 1 : Comparaison générale de la langue tchèque avec la langue française

(Source: http://www.unicaen.fr/typo_langues/consultation_langue.php?malang=fr)

1.1 Les inventaires phonémiques

Par rapport au système phonologique du français qui opère avec un nombre de voyelles variant de 13 à 16, selon les approches linguistiques adoptées, le tchèque possède 5 voyelles fondamentales. Celles-ci se manifestent sous les variantes "courte" ou "longue", leur durée vocalique ayant un rôle distinctif. A ces cinq voyelles fondamentales, courtes ou longues, viennent s'ajouter trois diphtongues.

Le système consonantique tchèque est par contre plus varié: 28 consonnes (contre 21 en français, dont 3 semi-consonnes).

Les deux systèmes phonémiques peuvent être résumés ainsi:

systèmes phonologiques	tchèque	français
nombre de voyelles	10 (5 "courtes" + 5 "longues") La durée vocalique est phonologique.	13 - 16
nombre de diphtongues	3 (dont 1 d'origine tchèque)	0
nombre de consonnes	28	18
nombre de semi-consonnes	0	3

Tableau 2 : Inventaire phonémique du tchèque et du français

L'articulation dans les deux langues est très différente, celle du français étant nettement plus marquée par la tension des muscles de l'appareil articulatoire que celle du tchèque. (Dohalská & Schulzová, 2003)

1.2 *Le niveau prosodique*

L'unité de perception et de production est la même, la syllabe, mais alors que la syllabation en français est influencée par le maintien ou la chute du *e* caduc ou instable et la contraction de certaines syllabes, en tchèque, la perte d'une syllabe serait perçue comme une déformation évidente. Par ailleurs, les mots lexicaux gardent leurs structures syllabiques dans les groupes rythmiques, autrement dit le processus de la resyllabation n'existe pas en tchèque.

Par contre les structures syllabiques de type CC, CCC ou CCCC existent et sont dues au caractère des consonnes syllabiques /r/ et /l/ (dans quelques cas rarissimes /m/ et /n/) et aux règles phonotactiques du tchèque.

Les deux langues sont des langues qui ont un accent fixe: la syllabe porteuse de l'accent principal est bien définie mais, alors qu'en tchèque c'est la première syllabe du "mot lexical", en français c'est la dernière syllabe du "mot phonétique" (Wioland, 2005: 28).

2. **Problèmes phonético/phonologiques et phonographématiques des apprenants tchèques**

Pendant les premières années d'apprentissage, la plupart des difficultés sont issues des multiples interférences entre les deux langues. L'impact d'une langue étrangère apprise avant le français ne sera pas traité ici car il n'y a pas encore, à notre connaissance, de résultats pour les recherches entreprises sur ce sujet. Néanmoins, nous avons constaté dans la pratique combien les interférences avec l'anglais étaient également fréquentes, notamment lorsqu'il s'agit du lexique commun aux deux langues et de sa graphie, comme *indépendant/independent*, par exemple.

Mais les problèmes de perception et, par conséquent, de production sont aussi posés par les particularités même du système phonologique français que nous avons présentées: d'une part le nombre et la variété des segments vocaliques du français, d'autre part sa structure syllabique qui estompe les limites entre les mots orthographiques font que les apprenants ne reconnaissent plus un mot familier appris isolément dès lors qu'il est intégré dans un groupe de sens ou une phrase complexe.

Enfin l'une des difficultés qui se présentent le plus souvent provient d'une graphie identique dans les deux langues pour des sons différents: alors que la graphie <e> est toujours prononcée [ɛ] en tchèque, nous savons combien

de sons elle peut représenter en français suivant son entourage. Non prononcée à la fin des mots, elle est cependant un indicateur important pour la prononciation de la consonne qui la précède.

En français, les sons entrent en contact dans la chaîne parlée et subissent de multiples influences provenant notamment des contractions, du *e instable*, des liaisons et des élisions, formant ainsi les groupes rythmiques (les groupes de sens). Pour les étrangers, ces groupes rendent plus difficile la perception des mots lexicaux et, de ce fait, le décodage du message lorsque leur oreille n'est pas entraînée à la perception de ces groupes.

3. La graphie et le son en interface

3.1 La réalité scolaire

Si nous considérons les descripteurs des niveaux B1-B2 du CECRL, les deux niveaux où se situent les étudiants qui obtiennent la *maturita* (le baccalauréat tchèque) qui leur permet de s'inscrire aux examens d'entrée des facultés, nous constatons combien les objectifs à atteindre sont exigeants par rapport à la réalité scolaire tchèque. Dans la partie du chapitre 5 intitulée *Les compétences de l'apprenant/utilisateur*, tableau *Maîtrise du système phonologique* (p. 92) nous lisons:

B2	A acquis une prononciation et une intonation claires et naturelles.
B1	La prononciation est clairement intelligible même si un accent étranger est quelquefois perceptible et si des erreurs de prononciation proviennent occasionnellement.

(source: http://www.coe.int/t/dg4/linguistic/source/framework_FR.pdf)

Nous estimons que, pour répondre à ces objectifs, la formation doit inclure un entraînement systématique des faits phonético/phonologiques.

Le développement du langage dans la classe de langue étrangère nécessite une réelle interactivité entre les apprenants et les formateurs. L'oreille de l'apprenant perçoit très vite les caractéristiques prosodiques de la langue: l'intonation, le rythme et les accentuations, mais il est nécessaire d'attirer immédiatement l'attention sur les différences essentielles entre la langue maternelle, la première langue étrangère en cours d'apprentissage (l'anglais en République tchèque) et la nouvelle langue étrangère (le français ou l'allemand pour les élèves tchèques). En effet, sans cette prise de conscience de la part de l'apprenant, jeune adolescent, adolescent ou adulte, le schéma phonologique de sa langue maternelle va être transféré dans la langue étrangère et ce phénomène d'interférence va créer un barrage à la compréhension puis à l'expression orale tout d'abord, au passage à l'écrit ensuite. Toutefois, nous croyons important de distinguer parmi les phénomènes phonologiques, comme le faisait déjà Walter (1977:

149-150), ce qui doit être considéré comme essentiel de ce qui est simplement contingent, comme l'intonation qui varie d'une région à une autre et d'un locuteur à un autre suivant sa personnalité ou son humeur du moment.

L'apprentissage de l'écrit est étroitement lié aux capacités métaphonologiques de l'apprenant et ce que dit Egaud au sujet des jeunes élèves souffrant de troubles spécifiques du langage oral et écrit est tout aussi valable pour les apprenants d'une langue autre que la langue maternelle:

"D'une part la conscience phonologique est nécessaire pour apprendre à lire et à écrire, d'autre part l'apprentissage de la lecture et de l'écriture améliore rapidement en retour la conscience phonémique." (Egaud, 2001)

Les meilleures conditions d'apprentissage de cette langue restent celles qui reproduisent les grandes étapes de l'apprentissage de la langue maternelle par le jeune enfant et l'on constate un nombre important d'interférences et de déficiences, à l'oral comme à l'écrit, lorsque ces étapes ne sont pas respectées. (Egaud, 2001, Hagège, 1996)

3.2 Du niveau secondaire au supérieur

L'évaluation de la conscience phonémique de nos apprenants est réalisée au début de leur première année d'études à l'université. Les étudiants proviennent d'établissements différents et ont été admis après un examen d'entrée qui détermine s'ils ont les connaissances suffisantes pour suivre les cours en français.

Cette première année est considérée comme une "mise à niveau" afin qu'ils aient les compétences requises pour poursuivre leurs études dans la filière qu'ils ont choisie. Puisque nous possédons les résultats obtenus à l'examen d'entrée (épreuves écrites et orales, niveau B1-B2 du CECRL) nous pouvons évaluer en fin d'année le chemin parcouru et recommander les points sur lesquels ils devront porter leur attention l'année suivante. Il s'agit, nous le voyons, d'une évaluation formative destinée à guider l'étudiant dans la suite de sa formation.

4. Notre expérience

Nous décrivons une expérience effectuée à la fin du mois d'avril 2011. Afin d'étudier, entre autres, l'influence des phénomènes phoniques sur l'écriture, nous avons opté pour la technique de la dictée. Cela nous permettait d'observer le niveau de l'interlangue des étudiants qui débutent leur formation pour devenir enseignant(e) de FLE et le type des erreurs persistantes.

4.1 *Données de départ*

Le texte de la dictée est d'un niveau B1 du CECRL pour le lexique et les structures grammaticales: aucun mot inconnu, plusieurs pronoms personnels, présent, passé composé, plus-que-parfait, plusieurs phrases complexes.

4.2 *Protocole du test*

<i>Objectifs</i>	Evaluer la capacité de perception orale et la compétence orthographique d'un groupe cible d'étudiant(e)s tchèques
<i>Hypothèse</i>	Les apprenants ayant déjà rencontré tous les éléments linguistiques repérés dans le texte, les erreurs proviendront de la subsistance d'une déficience dans le processus d'apprentissage / enseignement de la compréhension d'un message oral et de sa transcription écrite.
<i>Méthode</i>	Dictée (lecture à haute voix)
<i>Corpus</i>	Extrait de <i>Exercices de conversation et de diction pour étudiants américains</i> . E. Ionesco, In <i>Théâtre V</i> .

JM: Bonjour, Monsieur. Vous ne me demandez pas d'où je viens?

Ph: Je ne vous le demande pas car je le sais. Vous venez de Paris où vous avez passé une partie de vos vacances.

JM: Comment savez-vous que j'ai été à Paris?

Ph: C'est vous-même qui me l'avez dit, hier soir, quand je vous ai rencontré à la gare.

JM: Excusez-moi, j'avais oublié.

Ph: Au moins, est-ce que vous avez appris le français?

JM: Non, je n'ai pas pu l'apprendre. Les Parisiens parlent très mal le français. Ils le font sans doute exprès, car ils doivent connaître leur langue.

[...]

JM: Je n'ai pas fait grand-chose et je n'ai rien entendu d'intéressant puisque je n'entends rien, mais j'ai vu de très belles choses.

NB: Dans *A l'écoute des sons – Les voyelles*, page 64, Thérèse Pagniez-Delbart a choisi ce texte de Ionesco comme exercice de perception et de discrimination du son [ə] muet.

Lectrice de la dictée Enseignante universitaire d'origine française

Consigne (orale) "Je vais vous dicter un dialogue entre deux personnes, Jean-Marie (que vous écrirez JM) et Philippe (que vous écrirez Ph). Les signes de ponctuation vous seront dictés. Le texte sera lu une fois avant puis une fois après avoir été dicté. Vous n'aurez aucun temps de relecture ensuite."

Auditeurs et scripteurs 20 étudiant(e)s tchèques en première année de licence de la Faculté de Pédagogie (futurs professeurs de FLE potentiels).

Condition du test Test effectué dans une salle bien insonorisée

4.3 Les résultats et leur interprétation

Notre hypothèse de départ s'est trouvée confirmée: en dehors des erreurs "d'inattention" pour certains accords, les erreurs qui changeaient le sens du message provenaient d'une mauvaise discrimination phonologique et d'un manque de réflexion sur le sens de la phrase. Nous indiquons en gris les graphies qui ont été incorrectement transcrites par les étudiants:

JM : Bonjour, Monsieur. Vous ne me demandez pas d'où je viens ?

Ph : Je ne vous le demande pas car je le sais. Vous venez de Paris où vous avez passé une partie de vos vacances.

JM : Comment savez-vous que j'ai été à Paris ?

Ph : C'est vous-même qui me l'avez dit, hier soir, quand je vous ai rencontré à la gare.

JM : Excusez-moi, j'avais oublié.

Ph : Au moins, est-ce que vous avez appris le français ?

JM : Non, je n'ai pas pu l'apprendre. Les Parisiens parlent très mal le français. Ils le font sans doute exprès, car ils doivent connaître leur langue.

[...]

JM : Je n'ai pas fait grand chose et je n'ai rien entendu d'intéressant puisque je n'entends rien, mais j'ai vu de très belles choses.

- graphie incorrecte du son [e]: -ez/é transcrits e par 8 étudiants
- effet d'homophonie: d'où (écrit *doux* par 2 étudiants) / où (écrit *ou* par 3 étudiants)
- e muet non transcrit dans *partie* (3 étudiants)
- non perception de la nasale: *Au moins* (transcrit *Oh, moi* ou *au mois* par 8 étudiants)
- enchainement vocalique non détecté: *j'ai été* (écrit *j'étais* par 5 étudiants)

- confusion phonologique entre *rien* et *bien* (par 3 étudiants)
- graphie fautive dans *Je n'ai pas fait grand chose* (écrit *grande* par 5 étudiants) qui révèle que ces scripteurs ont fait usage de leur connaissance lexicale (*chose* mot féminin) pour compenser une perception orale déficiente
- non perception de la nasale: *entendu/entends* (écrit *attendu/attends* par 4 étudiants)
- perception déficiente: *de* (écrit *des* par la moitié des étudiants)

Ces résultats nous amènent à recommander un travail en amont, auprès des enseignants en activité dans les établissements primaires et secondaires et des futurs enseignants, pour qu'ils adoptent certaines démarches qui favoriseront la perception de la phonologie du français par leurs élèves.

"L'apprentissage devient celui de la perception, et le rôle du formateur est d'élargir le champ perceptif de l'apprenant." (Trocme-Fabre, 1987: 111)

A cette fin, nous proposons quelques principes qui pourraient être appliqués dans les cours de FLE:

- Parler et écrire, c'est d'abord écouter. Celui qui perçoit mal les sons les prononce mal et écrit incorrectement les messages, qu'ils soient dictés ou rédigés personnellement.
- Donner la préséance à l'oral sur l'écrit et opérer constamment un aller-retour entre la graphie et le son.
- Eviter d'apprendre à lire et/ou à prononcer des mots isolés, l'apprenant doit acquérir une prononciation syllabique correcte par l'articulation des groupes rythmiques. Prendre en compte le *e instable* et son rôle dans la chaîne prosodique.
- Entraîner aussi bien les faits prosodiques que phonématiques. En effet, ce sont les faits prosodiques qui influencent le caractère des segments dans le flux de la parole. De mauvais traits prosodiques (intonation, rythme et pauses) peuvent non seulement perturber la communication, mais aussi entraîner une interprétation fautive d'un message qui serait par ailleurs bien construit.

La mise en place de ces principes dès le début de l'apprentissage permettrait d'éviter les ancrages défectueux et les absences de repères entre le son et la graphie. L'acquisition systématique des compétences phonético / phonologiques favoriserait la compréhension des messages oraux et écrits et une expression écrite correcte.

Conclusion

Nous sommes convaincues de la pertinence de lier la phonie à la graphie et de développer une méthodologie qui favoriserait l'apprentissage des graphèmes par des phonèmes et vice versa (Pétrissans, 2010, Léon et al., 2009, Wioland, 1994). Des manuels comme *Plaisir des sons* (Kaneman-Pougatch & Pedoya-Guimbretière, 1989) et *A l'écoute des sons* (Pagniez-Delbart, 1990 et 1992) ont ouvert la voie dans ce sens.

Il ne reste dès lors qu'à souhaiter que les formations initiale et continue en République tchèque prennent en compte l'importance de cette association phonologie-graphie dans le processus de l'apprentissage/enseignement du FLE. Le nouveau manuel *Le français Entre Nous* destiné aux jeunes adolescents des établissements primaires et secondaires tchèques propose cette approche depuis une année (Nováková, Kolmanová, Geffroy Konštacký & Tábořská, 2009 et 2010). Les débuts de l'apprentissage sont exclusivement audio-oraux, le passage à l'écrit se faisant progressivement après plusieurs semaines. Après avoir appris à discriminer les nouveaux sons et après avoir réalisé naturellement la (re)syllabation du français par la reproduction de courtes formes telles que les comptines et de microdialogues, l'apprenant tchèque est moins perturbé par la découverte de la forme écrite. A partir du niveau 2, le côté sonore est mis explicitement en relation avec le code écrit. A travers diverses activités, l'apprenant est amené à la compréhension de cette relation. Il découvre alors qu'elle est plus transparente et régulière qu'en anglais (qu'il a considéré jusqu'alors comme une langue plus facile). Les nouvelles sonorités sont toujours introduites dans le contexte d'un "mot phonétique", ce qui habitue les apprenants à mieux reconnaître les "mots lexicaux" dans la chaîne parlée et à écrire ainsi correctement. Les premiers rapports des enseignants qui ont introduit cette démarche auprès de leurs élèves sont très positifs, il est toutefois encore trop tôt pour mesurer la portée des résultats au-delà de plusieurs années.

Nous recommandons les approches pédagogiques qui prennent en considération la langue maternelle des apprenants (Catach, 1994) et la découverte systématique des relations entre les graphies et les sons du français, parce qu'elles nous semblent le mieux répondre aux besoins de nos apprenants en favorisant leur prise de conscience des phénomènes particuliers à chacune des deux langues et, dans le cas qui nous intéresse ici, en leur donnant accès à une meilleure maîtrise de leur compétence orthographique.

Note: Nous appliquons l'orthographe recommandée par l'Académie française.

Bibliographie

- Catach, N. (1994): Le rôle de l'écrit dans l'acquisition d'une langue étrangère, ou l'utilisation des compétences. In: L'écrit en Français Langue Etrangère - Réflexions et propositions. Strasbourg (Presses Universitaires), 121-129.
- Dohalská, M. & Schulzová, O. (2003): Fonetika francouzštiny. Praha (UK, Karolinum).
- Egaud, C. (2004): Les troubles spécifiques du langage oral et écrit. Lyon (Centre régional de documentation pédagogique de l'Académie de Lyon).
- Geffroy Konšťacký, D. (2009): Un trésor de la langue française: l'orthographe moderne recommandée. Cahiers de Didactique du FLE dans les pays slaves, CADIFLESLAVES. Nitra, Slovaquie (Slovenská asociácia učiteľov francúzštiny).
- Hagège, C. (1996): L'enfant aux deux langues. Paris (Odile Jacob).
- Kaneman-Pougatch & M. Pedoya-Guimbretière, E. (1989): Plaisir des Sons. Paris (Hatier-Didier).
- Kross, I. (2011): L'orthographe reste une compétence stratégique. In: Le français dans le monde, Num. 373, 28-29. Paris (CLE International).
- Léon, P. et al. (2009): Phonétique du FLE: Prononciation de la lettre au son. Paris (Armand Colin).
- Nováková, S. (2005): La découverte des rapports entre la graphie et la prononciation en cours de FLE: le cas du "e muet". Plzeň (Sborník Fakulty pedagogické v Plzni, KRF, Cizí jazyky VII, ZČU), 109-115.
- Nováková, S., Kolmanová, J., Geffroy Konšťacký, D. & Táborská, J. (2009): Le français ENTRE NOUS 1. Plzeň (Fraus).
- Nováková, S., Kolmanová, J., Geffroy Konšťacký, D. & Táborská, J. (2010): Le français ENTRE NOUS 2. Plzeň (Fraus).
- Pagniez-Delbart, T. (1990): A l'écoute des sons. Les voyelles. Paris (CLE International).
- Pétrissans, J., (2010): Orthographe – Lire et entendre en même temps. In: Le français dans le monde, 371, 32-33. Paris (CLE International).
- Trocme-Fabre, H. (1987): Je pense donc je suis. Paris (Les Editions d'Organisation).
- Walter, H. (1977): La phonologie du français. Paris (PUF).
- Wioland, F. (1994): Le rôle de l'écrit en didactique de la prononciation du Français Langue Etrangère. In: L'écrit en Français Langue Etrangère – Réflexions et propositions. Strasbourg (Presses universitaires), 77-79.
- Wioland, F. (2005): La vie sociale des sons du français. Paris (L'Harmattan).

Sitographie

- http://www.unicaen.fr/typo_langues/consultation_langue.php?malang=fr (10. 4. 2010)
- http://www.coe.int/t/dg4/linguistic/source/framework_FR.pdf (15. 6. 2011)

OrthoFonic : un projet de didacticiel pour l'apprentissage de l'orthographe française

Claire FONDET & Fabrice JEJCIC

UMR 8589, CNRS-LAMOP, Paris 1 Panthéon Sorbonne

Histoire des systèmes graphiques du français et de ses variétés

Laboratoire de médiévistique occidentale de Paris

OrthoFonic is a lexical database of 5,500 words and 8,500 forms which is based on the *Listes orthographiques de base* (Basic Spelling Lists, 1984, 4000 forms). This list takes into account the revised spellings of the “Rectifications orthographiques” of 1990. To these Spelling Lists, which are in turn taken essentially from the *Frequency Dictionary of French Words* by A. Juilland *et al.* (1970), have been added the corresponding *Alfonic* phonological transcriptions and the IPA phonetic transcriptions as used in the *Petit Robert*. Thus, *OrthoFonic* could become the motor for a computerised teaching program constructed around the orthographical, phonological, phonetic and grammatical information associated with the most frequent lexical items of written French.

Creating a software program based on this high frequency lexis would make it possible to do away with forms that are rare or unused, and could encourage the acquisition of the written forms that are the most frequent in actual use.

Introduction

Chacun sait que la plupart des erreurs orthographiques, chez l'enfant comme chez l'adulte, sont commises sur les mots les plus ordinaires et les plus fréquents, plutôt que sur les mots rares (Ters, 1973: 60), cf. les exercices grammaticaux habituels concernant les homonymes monosyllabiques tels que: *a/à*, *ou/où*, etc. Ces observations plaident, avec évidence, en faveur d'une didactique de l'acquisition de l'écrit construite sur les formes les plus récurrentes de la langue.

Le projet *OrthoFonic* (combinaison de *orthographe* et *Alfonic*) ambitionne de renouveler l'enseignement du français courant en partant de formes vedettes et de formes fléchies classées selon un critère de fréquence. Les deux partenaires, le CNRS-LAMOP et l'Association RAPHAEL, recherchent actuellement un financement auprès de la région Île-de-France dans le cadre d'un PICRI¹, pour la mise au point d'un didacticiel d'apprentissage de l'écrit et de l'oral du français d'usage courant.

¹ RAPHAEL = Recherche pour l'Application de la PHonologie aux Apprentissages de l'Écriture et de la Lecture, PICRI = Partenariats Institutions-Citoyens pour la Recherche et l'Innovation. Le financement serait destiné à la mise au point informatique du didacticiel.

Ce projet pourrait intéresser à la fois le public scolaire (niveau élémentaire et début du secondaire), les personnes en difficulté (enfants et adultes éprouvant des gênes langagières, qu'elles soient orales ou écrites), les enfants scolarisés nouvellement arrivés en France (français langue seconde), les personnes en situation d'apprentissage du français langue étrangère (oral et écrit) et les migrants en situation d'alphabétisation dans les pays francophones².

1. Un vocabulaire scientifiquement construit

Les *Listes orthographiques de base* (1620 mots, 4000 formes)³ étaient, d'une certaine manière, et dit en toute modestie, en avance sur le temps quant à leur utilisation. À l'époque de leur élaboration sous la direction de Nina Catach, au début des années quatre-vingt et publiées en 1984, il n'existait pas encore de moyens techniques pour les utiliser à l'école autrement que sur un support papier. L'informatique était centralisée, les programmes et les données fonctionnaient avec des cartes perforées et, en dehors des laboratoires scientifiques, leur exploitation informatique n'était pas possible. Chaque tri, des quelque 4000 cartes comportant les mots de ces listes, coûtait l'équivalent de 2,60 € et l'on recevait le lendemain à Paris les tris imprimés des résultats du supercalculateur d'Orsay qui étaient livrés par navette. Aujourd'hui, la base complète, avec deux fois plus de données et trois fois plus de champs, nous livre un tri, le temps de taper sur la touche *entrée* du clavier d'ordinateur. Ce changement technologique et la banalisation de l'usage du micro-ordinateur ont motivé l'élaboration du projet actuel.

La base de données lexicales *OrthoFonic* (5500 mots, 8500 formes), fondée pour une grande part sur les *Listes orthographiques de base* (version 1994: 5000 mots, 8000 formes), a la particularité de contenir les formes fléchies les plus fréquentes et les mots les plus fréquents (*a*, *ai*, *avait* et *ont* sont les formes fléchies les plus fréquentes du mot *avoir*, forme vedette du verbe, ou lemme, au sens d'entrée de dictionnaire). D'autre part les *Listes orthographiques de base* (LOB), dont le *Frequency Dictionary of French Words* (FDFW), 5082 mots (Juilland et al., 1970), avait servi de fondement, avaient été complétées par les principales listes de vocabulaire du français: *L'élaboration du français élémentaire*, 1063 mots (Gougenheim et al. 1956, 1971), *Le vocabulaire orthographique de base*, 7691 mots (Ters et al., 1968), *L'échelle Dubois-Buyse*, 3725 mots (Ters et al, 1969) et, enfin, le *Dictionnaire des fréquences du Trésor de la langue française*, 71415 mots

² La didactique du français langue seconde (FLS) se distingue de celle du français langue étrangère (FLE) et du français langue maternelle (FLM). Sans entrer dans le débat de cette classification, disons simplement que la didactique doit être adaptée à la situation.

³ Catach N., Jejcic F. et équipe HESO (1984): Les listes orthographiques de base du français, Paris, F. Nathan, 160 p.

(document interne du CNRS sous forme de volumes polycopiés). Toutes ces listes, construites à partir d'enquêtes ou de corpus informatisés, font apparaître une grande concentration du vocabulaire d'usage très courant: elles ont en commun la quasi-totalité des 1280 premiers mots du *Trésor de la langue française* dont la fréquence, rapportée au nombre d'occurrences, est supérieure à 5000 (*LOB*, 1984: 17).

2. Un vocabulaire performant, de haute fréquence d'usage

Ce travail de synthèse sur les listes de fréquence assure un haut rendement au vocabulaire sélectionné. D'après nos études, dont les résultats sont comparables à la plupart des autres études de fréquence du vocabulaire, nous obtenons, pour les 5082 mots du *FDFW*, les chiffres suivants:

- les 10 premiers mots couvrent près de 35 % de la fréquence d'usage⁴;
- les 100 premiers mots couvrent près de 68 % de la fréquence d'usage;
- les 1000 premiers mots couvrent près de 89 % de la fréquence d'usage.

Parallèlement à cette observation, le *Dictionnaire des fréquences du Trésor de la langue française* fait apparaître que les 50 premiers mots, auxquels on ajoute les formes des verbes *être* et *avoir*, représentent plus de 35 millions d'occurrences, soit 50% de la fréquence (*LOB*, 1984: 16).

Cette "performance lexicale" nous a conduits à penser que nous sommes ici en présence d'une forme de "noyau" de la langue qu'il convient de privilégier dans l'enseignement.

3. Un vocabulaire conforme à la norme orthographique à enseigner

Le lexique intègre les *Rectifications orthographiques de 1990*, publiées dans les *Documents administratifs* du *Journal officiel* du 6 décembre 1990. Ces rectifications, approuvées par l'Académie française, constituent, en principe, la nouvelle norme orthographique à enseigner (Bulletins Officiels de l'Éducation nationale: n° 5, du 12 avril 2007 et n° 6 du 28 août 2008). Dans la 9^e édition en cours de son *Dictionnaire*, l'Académie rappelle que les nouvelles variantes cohabitent aux côtés des variantes graphiques traditionnelles pratiquées par les générations antérieures et: "[qu'] aucune de deux graphies ne peut être tenue pour fautive".

⁴ Le coefficient d'usage est la mise en rapport complexe de la fréquence (nombre d'occurrences) et de la dispersion (distribution des occurrences suivant les différents genres de textes: romans, théâtre, essais, presse, textes scientifiques et techniques), cf. *FDFW*, *Introduction*, p. XXX-XXXIX (genres de textes) et p. LXVII-LXXI (usage).

Parmi les dictionnaires actuels, *Hachette*, depuis son édition de 2002, et le *Petit Larousse*, depuis son édition de 2012 (publiée en 2011), intègrent totalement les *Rectifications orthographiques de 1990*. Le *Petit Robert* les insère en partie au fur et à mesure des millésimes.

Voici quelques exemples de mots rectifiés présents dans *OrthoFonic*:

- *abime, accroître, aout, boîte, bruler, couter, croute, diner*, etc. (suppression de l'accent circonflexe sur *i* et *u*), idem pour *apparaître* et les autres verbes en *-aitre*;
- un *après-midi* / des *après-midis*, un *porte-avion* / des *porte-avions* (marque du pluriel systématisée sur le second élément du composé à trait d'union);
- *à priori* (francisation de l'emprunt latin avec accent grave sur la préposition);
- *assoir* (correction de l'anomalie graphique par suppression du *-e-* parasite);
- *cèdera* (accent grave conforme à la prononciation pour le verbe *céder* et les autres verbes du même type);
- *asséner, évènement* (alignement de l'accentuation sur la prononciation);
- francisation des emprunts (accentuation, soudure de certains composés et marques de nombre), qu'ils soient latins (*référendum, arborétums*), italiens (*spaghettis, à capella*), anglais (*weekend, cowboys*), ou appartiennent à d'autres langues encore.

Les *Rectifications orthographiques de 1990*, par les modifications apportées, représentent donc une rationalisation du système graphique.

4. Une structure lexicale élaborée pour la didactique

OrthoFonic présente une structure hiérarchisée composée de 6 listes complémentaires de vocabulaire (voir Fig. 1, ci-dessous).

Les listes LBØ (LB "zéro"), LB1, LB2 et LB3 – qui composent les *LOB* 8000 formes – fonctionnent sur des indices de fréquence décroissante dont la distribution est globalisée à l'intérieur de quatre classes de fréquences. Ainsi la liste LBØ contient-elle les formes lexicales les plus fréquentes et la liste LB3 les formes lexicales les moins fréquentes.

La liste LB4 comprend le vocabulaire le plus fréquent en usage dans la presse (Skupien-Dekens C., Kamber A. & Dubois M., 2010: Avant-propos) et un lexique de la vie pratique issu de méthodes d'alphabétisation. La liste LB4 ne fonctionne pas sur des indices de fréquence calculés, mais intègre un lexique pratique courant, absent des *LOB*.

La liste LB5 est prévue pour accueillir des lexiques personnels, des vocabulaires spécifiques aux utilisateurs ou aux groupes d'utilisateurs.

Cet ensemble de listes peut être représenté par le tableau suivant:

N° liste / nb formes	N° d'usage FDFW	Fréquence ou type
LB0 - 157 formes	inférieur à 100	supérieure à 1000
LB1 - 403 formes	entre 100 et 250	supérieure à 30
LB2 - 3459 formes	inférieur à 1280	supérieure ou = à 5
LB3 - 4000 formes environ	supérieur à 1280	inférieure à 5
	Lexiques complémentaires	Types vocabulaires
LB4 - 500 formes environ	lexiques de la vie pratique	méthodes d'alphabétisation vocabulaire de la presse
LB5 - nb de formes ouvert	lexique personnel / du groupe	vocabulaires des usagers

Fig. 1: Hiérarchie de la structure lexicale

4.1 La fréquence du vocabulaire, listes LB0 à LB3

La notion de fréquence d'usage est tout particulièrement productive pour la didactique. Si nous prenons pour exemple le verbe *être* qui, toutes formes conjuguées réunies, arrive au 4^e rang de la fréquence d'usage, ceci indique que ce verbe est le 4^e mot le plus courant. Ce mot vedette représente toute une série de formes verbales; ses formes conjuguées apparaissent dans les trois premières listes:

- *est*, *était* et *sont*, dont la fréquence est supérieure à 1000, font partie de LB0;
- *es*, *étaient*, *étions*, et encore quelques autres formes dont la fréquence est supérieure à 30, appartiennent à la liste LB1;
- *fussent*, *serions*, *soyez*, et encore quelques autres formes dont la fréquence est supérieure ou égale à 5, relèvent de la liste LB2.

Toutes ces formes fléchies ont comme dénominateur commun le numéro du rang d'usage qui est le 4.

4.2 Les lexiques complémentaires, liste LB4

La liste LB4, qui n'est pas construite sur des calculs de fréquence, est cependant élaborée à partir du vocabulaire le plus fréquent de la presse écrite et à partir de lexiques de méthodes d'alphabétisation. À ce jour, cette liste comprend 500 mots supplémentaires, permettant d'actualiser les premières listes LB0 à LB3.

Les tableaux, ci-dessous (Fig. 2 et 3), présentent:

- colonne 1: la forme orthographique;

- colonnes 2, 3 et 4: la nature grammaticale, le genre et le nombre;
- colonne 5: la provenance de la forme: code PR = lexique de la presse, code AL = méthodes d'alphabétisation;
- colonne 6: la transcription phonologique en Alfonic;
- colonne 7: la transcription phonétique en alphabet phonétique international (API). Pour ces deux dernières, cf. § 4.6.

1 - Orthographe	2 - NG	3 - Ge	4 - Nb	5 - L	6 - Alfonic	7 - API
bancaire	AJ	S		PR	bäcer	bãkɛɾ
banlieue	NO	F	S	PR	bälix	bãljø
bicyclette	NO	F	S	PR	bisiclet	bisiklɛt
bloquer	VE			PR	bloce	bloke
boursier	NO	M	S	PR	bwrsie	bursje
caméra	NO	F	S	PR	camera	kamera
canal	NO	M	S	PR	canal	kanal

1 - Orthographe	2 - NG	3 - Ge	4 - Nb	5 - L	6 - Alfonic	7 - API
biscotte	NO	F	S	AL	biscot	biskɔt
boisson	NO	F	S	AL	bwasö	bwasõ
boiter	VE			AL	bwate	bwate
bonbon	NO	M	S	AL	böbö	bõbõ
bougie	NO	F	S	AL	bwji	buʒi
bouillon	NO	M	S	AL	bwyö	buʒõ
boulangerie	NO	F	S	AL	bwläjxri	bulãʒɛɾi

Fig. 2 et 3: Liste LB4: extraits du vocabulaire le plus fréquent de la presse et extraits du vocabulaire de méthodes d'alphabétisation.

4.3 *Le lexique utilisateurs, la liste LB5 ouverte*

La liste LB5 est une partie ouverte destinée à accueillir le lexique personnel de l'utilisateur ou du groupe. Ainsi, par exemple, dans le cadre d'activités spécifiques de la classe, les élèves pourraient intégrer dans cette liste un lexique spécialisé comme celui des sciences de la nature, celui de la géographie ou d'une autre discipline dont les vocables n'appartiennent pas à la langue courante. Et, pourquoi pas aussi des mots du vocabulaire personnel de l'apprenant?

4.4 *Le codage orthographique*

Une des difficultés de l'orthographe française provient des lettres muettes et des liaisons obligatoires ou facultatives. C'est pourquoi, en 1984, on avait choisi un codage pour marquer les différents types de lettres

muettes⁵, et un autre codage pour indiquer les liaisons, lorsque la forme considérée est placée devant un mot à initiale vocalique. Il s'agit essentiellement des lettres finales *n, r, s, t* (LOB, 1984: 20-23)⁶. Le tableau ci-dessous illustre ce système et nous donne le nombre de formes relevant de chaque code:

Codage	Type	Exemple	Nb formes
• code 7	lettres lexicales	<i>doigt77</i>	1122
• code 8	lettres grammaticales	<i>aimées88</i>	3116
• code =	lettres susceptibles de liaison	<i>chers8=</i>	373

Le code 7 indique des lettres muettes de type lexical, le code 8 les lettres muettes de type grammatical et le signe = indique les liaisons potentielles (laissées à l'appréciation de l'enseignant). Dans l'exemple *doigt77*, où ni le -*g*- ni le -*t* ne se prononcent, les codages du -*g*- par 7 et du -*t* par 7 établissent un lien avec les mots de la même famille: *doigté* et l'étymologie latine *digitum*. Par contre, ces deux lettres sont prononcées dans la forme dérivée *digital*. Cela correspond à la fonction dérivative de ces graphèmes qu'il est intéressant de transmettre dans le cadre d'un apprentissage. Dans l'exemple *aimées88*, où les deux dernières lettres ne sont pas prononcées, le codage du -*e*-, marque de féminin par 8, et le codage du -*s*-, marque de pluriel par 8, renseignent sur la nature grammaticale de ces graphies. Car le -*e* est souvent une façon de marquer le féminin d'une forme, et le -*s* est très fréquemment la marque du pluriel des adjectifs et des noms, bien qu'il indique aussi la marque verbale de la 2^e personne du singulier. Dans l'exemple *chers8=*, les codes 8 et = signifient respectivement que l'adjectif est au masculin pluriel (code 8) et que le -*s* peut être prononcé devant un mot commençant par une voyelle, comme dans la séquence: *chers8= amis*. Ce système de codage est une aide précieuse à la compréhension et à la maîtrise d'une partie difficile du système grapho-phonique du français.

4.5 Le codage grammatical

Le codage grammatical concerne la nature grammaticale des formes (parties du discours), le genre, le nombre, et indique les mots vedettes (cf. § 1).

⁵ Les chiffres 7 et 8 utilisés dans ce codage sont liés au fait qu'à l'époque le codage informatique était limité aux lettres majuscules sans accent. Ainsi, afin de pouvoir traiter les diacritiques, le code 1 représentait l'accent aigu (ex.: E1=é), le code 2 l'accent grave, le code 3 l'accent circonflexe, le code 4 le tréma, le code 5 la cédille. De plus le chiffre 6 codait le nom propre (ex.: PARIS6).

⁶ La sélection de ces cas de liaison a reposé uniquement sur le sentiment linguistique des auteurs à l'époque.

- Codage du champ "nature grammaticale"; les codages suivants sont utilisés:

Code	Valeur	CJ	conjonction	NU	numéral
AJ	adjectif	IN	interjection	PN	pronom
AR	article	LO	locution	PS	préposition
AV	adverbe	NO	nom	VE	verbe

- Codage du champ "genre"; les codages suivants sont utilisés:

Code	Valeur	Remarque:
F	féminin	l'absence des codes de notation du genre, M ou F, indique qu'il s'agit d'une forme épïcène
M	masculin	

- Codage du champ "nombre"; les codages suivants sont utilisés:

Code	Valeur	Remarque:
S	singulier	l'absence des codes de notation du nombre, S ou P, indique qu'il s'agit d'une forme invariable
P	pluriel	

- Codage du champ "vedette"; le codage suivant est utilisé:

Code	Valeur	Remarque:
*	forme vedette	la présence de l'astérisque dans le champ indique une forme vedette, type entrée de dictionnaire (ex.: <i>*être</i>); l'absence de l'astérisque dans le champ indique une forme fléchie, réalisée dans le discours (ex.: <i>était</i> 8=)
	forme fléchie	

4.6 Les codages phoniques

Pour le versant oral de la langue, *OrthoFonic* présente deux codages phoniques: l'API et l'Alfonic. L'API fut inventé pour noter toutes les langues du monde, en 1888, par Paul Passy. Ce codage est plus particulièrement adapté à l'apprentissage du français langue étrangère. En revanche pour les élèves francophones, André Martinet, à la demande d'enseignants dans les années soixante-dix, mit au point le code phonologique dénommé Alfonic, selon les oppositions pertinentes du français oral. Le but était d'enseigner l'écrit à partir de l'oral. L'utilisation du clavier de la machine à écrire et la proximité avec la graphie du français ont été privilégiées. Ce code était considéré comme un outil que les élèves devaient s'approprier depuis la dernière année de Maternelle jusqu'au Cours Préparatoire. Dans la perspective de l'acquisition de l'orthographe, l'emploi de l'outil Alfonic qui est conçu comme une écriture de transition, s'est révélé très efficace, conduisant à la mise en place d'une conscience critique de la norme orthographique (Vial J. et Villard J., 1986: 1-2).

Peu d'écoles pratiquent l'Alfonic aujourd'hui, mais un didacticiel a été élaboré en application du projet européen *Je parle donc j'écris* (cédérom et

livret produits en 2005). Ce logiciel est accessible gratuitement (voir sites internet). L'Association RAPHAEL poursuit ce projet.

Certains caractères de l'Alfonic sont différents de ceux de l'API (quelques consonnes et la plupart des voyelles). Pour les consonnes il n'y a pas de différence de signes entre les deux codes. Inversement, certaines voyelles sont notées différemment. L'Alfonic suit la théorie de la phonologie présentée dans les *Éléments de linguistique générale* d'André Martinet. Sur le plan phonologique, il n'y a pas de différence entre les deux codes en ce qui concerne le système consonantique. La question se pose essentiellement pour la notation de certaines oppositions vocaliques (loi de position). Il s'agit des voyelles orales d'ouverture moyenne correspondant aux archigraphèmes É, Ê, EU, O.

(1) en Alfonic: /è/ ~ /e/, en API: [ɛ] ~ [e]; (2) en Alfonic: /o/ ~ /ô/, en API: [ɔ] ~ [o]; (3) en Alfonic /x/ ~ /x^/, en API: [œ] ~ [ø].

- 1) Opposition /è/ ~ /e/, en API [ɛ] ~ [e]: si nous prenons la forme verbale *prêtait* (707), celle-ci est transcrite ici en Alfonic: /pretè/ et en API: [pʁɛtɛ]. La différence réside ici dans la voyelle de la première syllabe. Car, du point de vue phonologique, il n'existe pas en français d'opposition /e/ ~ /è/ en syllabe intérieure, c'est-à-dire non finale, même si l'orthographe autorise une prononciation ouverte de la voyelle en question. On considère en effet que dans ce cas /e/ et /è/ sont en distribution complémentaire ([e] en syllabe ouverte, [ɛ] en syllabe fermée). Le fait de prononcer /è/ en syllabe ouverte est ici une variante libre, à la différence de *pré* / *prêt* (Alfonic /pre/ ~ /prè/, API [pʁɛ] ~ [pʁɛ]). On trouvera donc en API [ɛ] en toute position, contre /è/ seulement en finale absolue en Alfonic. Autres exemples: *perdu*, Alfonic \perdu\, API [pɛʁdy]; *paisible*, Alfonic \pezibl\, API [pɛzibl]; *prévu*, Alfonic \prevu\, API [pʁevy].
- 2) Opposition /o/ ~ /ô/, en API: [ɔ] ~ [o]. L'API transcrit partout la différence voyelle ouverte ~ voyelle fermée, tandis que l'Alfonic ne garde ici cette opposition qu'en syllabe finale fermée. Exemples: *notre* = Alfonic: /notr/ et API: [nɔʁ]; *le nôtre* = Alfonic /lx nôtr/ et API [lə nɔʁ]. Et par ailleurs *porter* donne en Alfonic \porte\ et en API [pɔʁte]; *poser* donne \poze\ en Alfonic et en API [poze].
- 3) Même chose pour l'opposition /x/ ~ /x^/, en API: [œ] ~ [ø]. L'Alfonic ne garde cette opposition que dans le cas de la voyelle en syllabe finale fermée. Cette opposition est bien représentée, car même s'il n'existe que de rares paires minimales (ex.: *jeune* ~ *jeûne*), le suffixe *-euse* est productif, ex.: *heureuse*.

Entre l'API et l'Alfonic, tous deux présents dans *OrthoFonic*, quelle notation faudra-t-il privilégier? Telle est la question qui se pose à l'enseignant. Cela dépendra de la situation pédagogique. L'API peut paraître plus "juste" pour noter les manières différentes de prononcer l'archigraphème EU dans *fleur*

et dans *feu*. L'expérience montre en effet que, dans les cas de neutralisation de voyelles, les enfants adoptent spontanément une démarche plus phonétique que phonologique⁷.

En fait, la différence entre les deux codes n'a pas été d'avantage accentuée. C'est ainsi que l'effacement du *schwa* intérieur et du *-e* final – ex.: *bib(e)ron*, *ball(e)* – ont été harmonisés dans les deux types de notation.

Quel que soit le code préféré, l'étude de la face phonique du vocabulaire fera prendre conscience de la variation dans l'oral et permettra de développer un esprit de tolérance à ce sujet. Certes, la base *OrthoFonic* peut paraître rigide, mais constitue un point de départ pour alimenter la réflexion.

Dans l'avenir, sur le plan technique, le codage en API pourrait permettre, via une interface appropriée, la reproduction sonore des vocables, utile en FLE et FLS, ainsi qu'en FLM, pour lever l'ambiguïté graphique des formes homographes hétérophones (Jejcic 1987: 37-39), ex.: *les poules du couvent couvent*.

5. La base de données *OrthoFonic*

Dans sa configuration actuelle, la base de données comprend environ 8500 enregistrements et 16 champs par enregistrement. Le tableau ci-après (Fig. 4) présente la base et les fonctions des divers champs.

Nom du champ	Exemple	Fonction
Nusage	0072	<i>numéro d'usage</i> , établi d'après le rang d'usage du <i>FDFW</i> . Plus ce chiffre est petit, plus la forme est fréquente. L'adjectif <i>petit</i> est au 72 ^e rang, ce qui signifie qu'il est le 72 ^e mot (au sens de lemme), le plus utilisé
Ved	*	la présence du signe "*" étoile, indique qu'il s'agit d'une forme vedette, ex.: * <i>petit7</i> , avec étoile, est une forme vedette; <i>petits78=</i> , sans étoile, est une forme fléchie de l'adjectif <i>petit7</i>
Formes_C	<i>petit7=</i>	<i>formes des mots avec lettres muettes et liaisons codées</i> : ici le code "7" indique une lettre muette de type lexical, dérivative ou non dérivative selon le cas, ex.: <i>petit7</i> donne <i>petite</i> ; le code "8" indique une lettre muette de type grammatical, pluriel, ex.: <i>petits78=</i> pluriel; le code "=" indique une liaison possible devant initiale vocalique, ex.: <i>petits78= enfants</i> , chaîne graphique prononcée [pətizãfã] avec la liaison
Alfonic	pxti	transcription phonologique en caractères Alfonic
Nat	AJ	<i>nature grammaticale</i> , 10 catégories: AJ: adjectif, AR: article, AV: adverbe, CJ: conjonction, IN: interjection, LO: locution,

⁷ Martinet, A. (1989: 12). Il s'agit de la neutralisation d'une voyelle dont l'ouverture ou la fermeture dépend du contexte comme dans le cas de *père*, autrefois prononcé [pɛrə]. Selon l'article cité, les enfants avaient tendance à noter *fête* en Alfonic: \fèt\ au lieu de \fet\.

		NO: nom, NU: numéral, PN: pronom, PS: préposition, VE: verbe
Ge	M	<i>genre</i> . Remarque: l'absence de code M ou F indique une forme épicène
Nb	S	<i>nombre</i> . Remarque: l'absence de code S ou P indique une forme invariable
Nl	0	<i>numéro de la Liste orthographique de base</i> : plus le numéro est petit, plus la fréquence de la forme est élevée. Les formes sont regroupées par "fourchette" de fréquence dans 4 listes: liste 0 [liste "zéro"] = mots très fréquents, liste 1 [liste "un"] = mots un peu moins fréquents, liste 2 = encore un peu moins fréquents, liste 3 = mots les moins fréquents; liste 4 = liste complémentaire ne fonctionnant pas sur des critères de fréquence, liste 5 = liste ouverte complétée par l'utilisateur
Sl	FD, DB, GG, HE, TL, AL, PR, _	<i>sous-liste</i> , indique la provenance de la forme ou son attestation dans une liste de vocabulaire: <i>FD</i> , la forme provient du <i>FDFW</i> ; <i>DB</i> , la forme provient de l' <i>Échelle Dubois-Buyse</i> ; <i>GG</i> , la forme provient du <i>Français fondamental</i> de G. Gougenheim; <i>HE</i> , la forme est codée par <i>HESO</i> (équipe de recherche du CNRS, 1973-1998), <i>TL</i> , la forme provient du <i>Dictionnaire de fréquence du Trésor de la langue française</i> (Institut National de la Langue Française-CNRS, aujourd'hui ATILF); <i>AL</i> vocables issus de méthodes d'alphabétisation; <i>PR</i> mots issus du vocabulaire fréquent de la presse; <i>absence de marque</i> , la forme provient de la liste ouverte par l'utilisateur
Formes	petit	<i>formes orthographiques</i> des mots (sans le codage des lettres muettes)
Inverse	titep	<i>formes inverses</i> des mots sans le codage des lettres muettes: les mots sont écrits "à l'envers", la dernière lettre du mot a le rang un et la première lettre du mot le dernier rang, ex.: <i>titep</i> est l'inverse de la forme <i>petit</i> - ceci permet de regrouper et trier des formes de mots qui ont des finales identiques
Inverse_C	7titep	<i>formes inverses</i> des mots avec le codage des lettres muettes: ex: <i>7titep</i> est l'inverse de la forme <i>petit7</i> - ceci permet de regrouper et trier soit des formes de mots qui ont des finales et des codes identiques (toutes les formes terminées par un -t codé 7), soit toutes les formes terminées par un code 7, indépendamment de la lettre qui précède le code
Syllab	pe tit	<i>syllabisation graphique de la forme</i> , selon les règles orthotypographiques de coupure des mots. La coupure syllabique est indiquée par un blanc graphique interne au mot: <i>pe_tit</i>
Phonetique	PàTI	<i>transcription phonétique</i> de la forme en code machine dont la correspondance en API est [pəti]
Syll_Phon	Pà/TI	<i>transcription phonétique</i> de la forme en code machine avec indication par un / "slash" de la coupure syllabique et dont la correspondance en API est [pə/ti]
Phon_Inv	ITàP	transcription phonétique inverse des formes, ex.: [itəp] est l'inverse de la forme [pəti], ceci permet de regrouper toutes les finales à prononciation identique

Fig. 4: Structure de la base de données *OrthoFonic*.

6. L'exploitation didactique de la base de données *OrthoFonic*

Les exemples qui suivent représentent quelques pistes d'utilisations diverses de la base sur les graphies des finales:

- les oppositions finales codées / finales non codées (§ 6.1.1)
- les finales codées de liaisons potentielles (§ 6.1.2)
- les finales codées de formes homophones (§ 6.1.3)
- les adjectifs épïcènes (§ 6.1.4)
- les noms féminins terminés par *-t* et les noms masculins terminés par *-e* (§ 6.1.5)

6.1 *Les recherches sur les graphies des finales*

Nous l'avons déjà mentionné plus haut: une grande partie des difficultés de l'orthographe provient des mots familiers et des lettres finales muettes. À ce dernier point en particulier correspondent des fonctionnements lexicaux et grammaticaux divers.

Une observation sur les graphies finales les plus fréquentes nous donne le tableau ci-dessous (Fig. 5), les graphies les plus fréquentes étant placées à gauche:

Nombre de finales par type de lettre								
-s	-r	-t	-n	-e	-x	-z	-d	-p
1600	1102	1053	463	414	142	87	68	11

Fig. 5: Recherche sur les lettres finales les plus fréquentes

Trois ensembles se distinguent nettement. Le premier, avec les finales *-s*, *-r* et *-t*, marque la prédominance du *-s*. La polyvalence du *-s* est distribuée entre la marque générale du pluriel des adjectifs et des noms, mais aussi la marque de seconde personne du singulier et de la première personne du pluriel du paradigme verbal, et souvent encore il peut être un marqueur graphique de l'adverbe et de la préposition, etc. Des remarques similaires s'appliquent au second ensemble, avec les finales *-n*, *-e*, et *-x*, et au troisième ensemble avec les finales en *-z*, *-d* et *-p*.

Ce type d'investigation sur les mêmes finales peut être poursuivi sur l'absence et/ou la présence du codage graphique des finales.

6.1.1 Les oppositions finales codées / finales non codées

L'absence de code en finale indique que la lettre est prononcée. Dans le cas contraire, elle est codée 7 ou 8, ce que montrent les exemples qui suivent:

-s est prononcé dans *autobus*, il est muet dans *abus7* où le code, de type lexical, indique une dérivation possible comme dans *abuser*, *abusif*, *abusivement*; il est aussi muet dans *abeilles8* où le code, de type grammatical, indique le pluriel du nom *abeilles*. Le tableau ci-dessous (Fig. 6) donne une vue synthétique de ces fonctionnements, leur nombre d'occurrences (abréviation: nb.) et quelques exemples.

finale	nb.	exemples	finale	nb.	exemples	finale	nb.	exemple
-s	12	<i>autobus</i>	-s7	156	<i>abus7</i>	-s8	1229	<i>abeilles8</i>
-r	329	<i>aboutir</i>	-r7	66	<i>acier7</i>	-r8	705	<i>chanter8</i>
-t	27	<i>brut</i>	-t7	641	<i>abondant7</i>	-t8	299	<i>fait8</i>
-n	446	<i>abandon</i>						
			-e7	122	<i>entrée7</i>	-e8	251	<i>jolie8</i>
						-es78	41	<i>pensées78</i>
			-x7	43	<i>boiteux7</i>	-x8	51	<i>animaux8</i>
-z	1	<i>gaz</i>	-z7	2	<i>nez7</i>	-z8	75	<i>aimez8</i>
-d	1	<i>sud</i>	-d7	60	<i>abord7</i>			
-p	2	<i>cap</i>	-p7	8	<i>beaucoup7</i>			

Fig. 6: Oppositions entre les lettres finales les plus fréquentes

Dans l'exemple *abandon*, le -n est censé être prononcé, mais en réalité il est le second élément du graphème *on*, notant la voyelle nasale. La finale -x, marque lexicale dérivative dans *boiteux7*, est une marque de pluriel dans *animaux8*. Contrairement au -s final, très récurrent, le -x final est une graphie moins fréquente mais présente sur des noms et des adjectifs très courants, dont la dérivation est systématiquement réalisée par l'alternance de -x en -s-, *boiteux7* / *boiteuse*, *heureux7* / *heureuse*, etc. La finale -z correspond au second élément graphique du graphème *ez*, marque systématique de la seconde personne du pluriel des verbes. Par contre, dans *nez*, *chez* et *rez-de-chaussée*, le z a été conservé pour des raisons de distinction, notamment d'homonymes pour *nez* (cf. *DHOF*: 707).

6.1.2 Les finales codées de liaisons potentielles

Les liaisons, devant un mot à initiale vocalique, paraissent évidentes pour quelques mots-outils d'usage très courant: *pas7=*, *deux7=*, *est8=*, *grand7=*, *tes8=*, etc., voir tableau (Fig. 7).

let.	nb.	exemples	let.	nb.	exemples	let.	nb.	exemple
-s=	14	<i>en bas= âge</i>	-s7=	21	<i>après7=</i>	-s8=	168	<i>aimables8=</i>
			-x7=	42	<i>couteux7=</i>	-x8=	6	<i>aux8=</i>
			-z7=	2	<i>assez7=</i>	-z8=	7	<i>allez8=</i>
-t=	19	<i>bout= à bout7</i>	-t7=	23	<i>avant7=</i>	-t8=	44	<i>allait8=</i>
-d=	2	<i>de pied= en cap</i>	-d7=	5	<i>grand7=</i>			
			-p7=	1	<i>trop7=</i>			
			-r7=	2	<i>dernier7=</i>			

Fig. 7: Locutions courantes et mots fréquents produisant la liaison

Certes ces liaisons ne sont pas à mettre toutes sur le même plan: seules deux d'entre elles sont obligatoires ici. L'économie de la base *OrthoFonic* ne tient pas compte de cette différence. Cependant, le codage de certaines liaisons moins évidentes permet une réflexion sur les registres de langue, ce qui est le cas de l'adjectif *couteux7=* dont la liaison suppose un style soutenu.

D'autre part, environ 80 locutions courantes, ou expressions figées (*de pied= en cap*), ont été intégrées aux listes. Leur intérêt est d'apporter à l'enseignement de l'orthographe des compléments d'information sur des liaisons internes anciennes qui souvent ne se font plus par ailleurs.

6.1.3 Les finales codées de formes homophones

La distinction apportée par le système de codage est une des clés montrant les diverses fonctions des marques morphologiques de l'orthographe. La requête qui demande à la base de lister tous les mots prononcés en [fɛ], nous donne à l'écran le résultat suivant:

FORMES-C	NAT	GE	NB	SL
fais8=	VE		S	1
fait7	VE	M	S	1
fait8=	VE		S	1
faits78	VE	M	P	2
fait7=	NO	M	S	1/2
faits78	NO	M	P	1
fait7	AJ	M	S	3
faits78	AJ	M	P	3

Fig. 8: Listage de tous les mots qui se prononcent en [fɛ]

Les mots listés nous montrent l'opposition entre les verbes, les noms et les adjectifs ainsi que l'opposition entre les morphogrammes lexicaux et les morphogrammes grammaticaux. Cette différenciation entre les codes 7 et 8 est illustrée par les occurrences d'homographes *fait7=* et *fait8=*. Dans le

premier exemple, le code 7 indique une lettre muette de type lexical qui est rattachée à la catégorie du nom masculin singulier, le code =, qui s'applique à la même lettre, signale que le -t final est prononcé devant initiale vocalique, comme dans la locution "prendre fait et cause". En revanche, dans le second exemple, le code 8, désigne une lettre muette de type grammatical qui dépend de la catégorie du verbe au singulier. Le code = signale une liaison facultative, ex.: "il fait encore beau". Il est possible d'analyser ainsi chacune des occurrences listées.

Il nous semble que ce processus de découverte des sous-systèmes graphiques de l'orthographe contribue à une acquisition solide des connaissances nécessaires à la maîtrise de l'écrit.

Parallèlement au travail d'analyse graphique, l'observation peut aussi porter sur la fréquence des formes relevant des différentes catégories grammaticales: les formes [fe] du verbe et du nom relèvent toutes des listes LB1 et LB2 (code SL, informations de la dernière colonne à droite), contrairement aux deux dernières occurrences qui appartiennent à la liste LB3.

6.1.4 Les adjectifs épïcènes

Si la conscience linguistique paraît bien établie sur la séparation des genres – féminin et masculin – pour les apprenants en FLM, une autre piste de réflexion peut être amorcée sur l'absence de marque de genre, lorsque la forme graphique se termine en -e. C'est le cas des adjectifs épïcènes dont voici quelques exemples.

FORMES-C	NAT	GE	NB	SL
difficile	AJ		S	2
difficiles8	AJ		P	2
pauvre	AJ		S	2
pauvres8=	AJ		P	2
facile	AJ		S	2
faciles8	AJ		P	2
nécessaire	AJ		S	2
nécessaires8	AJ		P	2
politique	AJ		S	3
politiques8	AJ		P	3

Fig. 9: Extraits du listage des adjectifs dont la colonne genre (GE) est vide

L'exercice de la commutation "une femme *pauvre* / un homme *pauvre*" valide l'absence de genre marqué. L'observation de la fréquence des formes indique que l'adjectif *politique*, liste LB3, est moins fréquent que les autres, liste LB2, ce que confirme aussi le sentiment linguistique de tout un chacun.

D'autre part, ces investigations peuvent être étendues à la sphère du nom. La notion de nom épïcène a facilité les changements récents liés à la féminisation des noms de métiers, titres, grades et fonctions (*FJTN*: 1999). Par exemple, *professeur* et *ministre* sont épïcènes phonétiquement parlant, mais graphiquement le genre féminin est indiqué par l'ajout d'un -e, marque de féminin, à *professeure* et d'une marque syntaxique par le déterminant féminin *une/la* devant le vocable *ministre*.

6.1.5 Les noms féminins terminés par -t et les noms masculins terminés par -e

On peut s'intéresser à un paradoxe: certains noms féminins sont terminés par -t, d'autres, masculins, le sont par -e. On constate que les premiers sont beaucoup moins nombreux que les seconds. On peut se poser encore d'autres questions (étymologie, difficulté de mémorisation du genre des noms masculins terminés par -e, par exemple le genre de *solde* qui a deux entrées dans le dictionnaire).

190	part7	698	plupart7	2335	minuit7
208	nuit7	1035	dent7		
345	mort7	1094	forêt7		

Fig. 10: Listes noms féminins terminés par -t

70	homme	678	verre	1250	ministère
268	compte	695	membre	1302	fleuve
271	exemple	702	théâtre	1344	capitaine
277	père	705	système	1503	mystère
300	livre	709	genre	1532	décembre
324	service	758	arbre	1563	kilomètre
329	maitre	787	terme	1586	programme
337	problème	789	pote	1644	musée
343	nombre	815	passage	1646	septembre
352	caractère	822	centre	1677	ange
354	être	859	commerce	1703	poste
356	ordre	884	crime	1781	comte
358	siècle	895	rêve	1795	souffle
361	peuple	899	ouvrage	1814	charme

Fig. 11: Extraits de la liste de noms masculins terminés par -e

Ce ne sont-là que quelques exemples de recherches que les apprenants, quels que soient leur niveau et leur âge, pourront effectuer selon tel ou tel critère. Autre possibilité: le maître ou le formateur proposera en vrac une liste de mots à classer sans autre consigne, pour amener l'élève à découvrir la structure de la langue. Si cet exercice est orienté vers une solution unique, la présence d'un intrus la rendra plus ludique. La base *OrthoFonic* est un outil dont chacun pourra découvrir les possibilités.

7. Les objectifs du projet

Le succès de la candidature PICRI serait le début d'un partenariat réunissant l'ensemble des acteurs pour la mise au point du didacticiel en question. Ce partenariat impliquerait, outre le CNRS-LAMOP et l'association RAPHAEL déjà cités, d'une part des étudiants inscrits en Master 2 de traitement automatique des langues, d'autre part des enseignants et leurs élèves. Les étudiants auraient à déployer une interface logicielle dans un langage le plus naturel possible pour les futurs utilisateurs. Avec le groupe de professeurs et directeurs d'écoles, serait établie une typologie des requêtes fondée sur leur expérience. On obtiendrait ainsi un outil de travail souple, capable de produire des sélections et des tris adaptés à des séquences didactiques déterminées. La mise au point du logiciel devrait être réalisée en relation avec les praticiens et leurs élèves. Pour ces derniers, la médiation de l'ordinateur, dans un travail en petits groupes, est un élément stimulant dont il faut tenir compte.

Ce programme concernerait des classes de CM1 et CM2 de l'école élémentaire et les classes de collège. Il bénéficierait notamment aux enfants en difficulté pour une meilleure intégration scolaire.

Le champ d'expérimentation, si l'on peut s'exprimer ainsi, concernerait en grande partie des classes situées en ZEP (zone d'éducation prioritaire), dans la commune de Villejuif, où la mixité sociale et les différences socioculturelles sont du plus grand intérêt pour une telle entreprise.

Par ailleurs, pourquoi ne pas envisager par la suite un autre volet destiné à des adultes (souvent parents d'enfants scolarisés), en exploitant les mêmes éléments de la base *OrthoFonic*, afin de développer un module d'alphabétisation?

Conclusion

Le caractère novateur de ce projet de logiciel interactif et dynamique repose sur:

- la haute fréquence d'usage du vocabulaire proposé;
- le système de codage grammatical des lexèmes et le système de codage orthographique des formes facilitant la maîtrise des diverses fonctions des lettres muettes de l'orthographe;
- l'appropriation systématique de la relation oral - écrit.

OrthoFonic vise de la sorte à renouveler les modalités d'acquisition par une approche multidimensionnelle montrant les relations entre système graphique et système phonique, mais aussi les relations à l'intérieur de l'écrit entre les divers sous-systèmes de l'orthographe.

À terme, nous espérons disposer gratuitement d'un tel logiciel, qui favoriserait la démarche heuristique et valoriserait les connaissances de l'apprenant au fur et à mesure que ce dernier consolide ses acquisitions.

Abréviations et sigles

API = Alphabet Phonétique International

CNRS = Centre National de la Recherche Scientifique

DHOF = Dictionnaire Historique de l'Orthographe Française (cf. Bibliographie: Catach, N. dir.)

FDWF = Frequency Dictionary of French Words (cf. Bibliographie: Juilland, A.)

FLE = Français Langue Étrangère

FLM = Français Langue Maternelle

FLS = Français Langue Seconde

FJTN = Femme, j'écris ton nom... (cf. Bibliographie: Becquer, A.)

LAMOP = Laboratoire de Médiévistique Occidentale de Paris

LOB = Listes Orthographiques de Base (cf. Bibliographie: Catach, N., Jejcic, F. & équipe HESO)

PICRI = Partenariats Institutions-Citoyens pour la Recherche et l'Innovation

RAPHAEL = Recherche pour l'Application de la PHonologie aux Apprentissages de l'Écriture et de la Lecture

TLF = Trésor de la Langue Française

UMR = Unité Mixte de Recherche

Bibliographie de l'article

Association Phonétique Internationale, Le maître phonétique, revue fondée par Paul Passy et éditée de 1904 à 1967.

Becquer, A., Cerquiglini, B., Cholewka, N., Coutier, M., Frécher, J. & Mathieu, M.-J. (1999): *Femme, j'écris ton nom... Guide d'aide à la féminisation des noms de métiers, titres, grades et fonctions*. Paris (CNRS-INaLF, La Documentation française).

Bulletins Officiels de l'Éducation nationale: n° 5, du 12 avril 2007 et n° 6 du 28 août 2008.

Catach, N. dir. (1995): *Dictionnaire historique de l'orthographe française*. Paris (Larousse).

Catach, N., Jejcic F. & équipe HESO (1984): *Les listes orthographiques de base du français*. Paris (Nathan).

Gougenheim, G., Michéa, R., Rivenc, P. & Sauvageot, A. (1956): *L'élaboration du français élémentaire*. Paris (Didier).

Gougenheim, G. (1971): *Le français élémentaire*. In: *Introduction du Grand Larousse de la langue française*, t. 1, p. LXXXII-XC. Paris (Larousse).

Jejcic, F. (1987): *Pour un programme minimal de phonétisation automatique du français*. In: *Liaisons-HESO*, 15, 33-59.

Juilland, A. et al. (1970): *Frequency Dictionary of French Words*. The Hague (Mouton).

Martinet, A. (1960): *Éléments de linguistique générale*. Paris (Armand Colin).

Martinet, A. & J. (1980): *Dictionnaire de l'orthographe Alfonic*. Paris (SELAF, Peeters-France).

- Martinet, A. & J., Villard, J., Boyer, D. & Dominici, G. (1983): Vers l'écrit avec Alfonic. Paris (réédition 2006, Association RAPHAEL).
- Martinet, A. (1989): Alfonic au plus près de l'orthographe. In: Liaison Alfonic, fascicule 2, 7-12.
- Les Rectifications de l'orthographe (1990): Journal Officiel de la République française, Documents administratifs, 100 (06.12.1990).
- Skupien-Dekens, C., Kamber, A. & Dubois, M. (2010): Manuel d'orthographe pour le français contemporain. Institut de langue et civilisation françaises (Université de Neuchâtel).
- Ters, F., Mayer, G. & Reichenbach, D. (1968): Vocabulaire orthographique de base. Neuchâtel (Messeiller), Paris (OCDL).
- Ters, F. (1973): Orthographe et vérités. Paris (ESF).
- Ters, F., Mayer G. & Reichenbach D. (1975): L'échelle Dubois-Buyse d'orthographe usuelle française. Neuchâtel (OCDL, Messelmer) [3e éd.].
- Trésor de la langue française, Dictionnaire de la langue du XIXe et du XXe siècle (1789-1960), 1971-1994, publié sous la direction de Paul Imbs puis de Bernard Quemada. Paris (CNRS-Klincksieck-Gallimard).
- Vial, J. & Villard, J. (1986): À l'entrée au cours élémentaire. In: Liaison Alfonic, fascicule 1, 1-2.

Bibliographie complémentaire

- Baddeley, S. & Jejcic, F. (2010): Autour des rectifications de l'orthographe, aspects historiques et actuels. De la "droicte escripture" aux Rectifications de 1990. In: Rectification à l'oral et à l'écrit, Paris (coll. L'Homme dans la langue, Ophrys), 191-206.
- Biedermann-Pasques, L. & Jejcic, F., eds. (2006): Les rectifications orthographiques de 1990. Analyses des pratiques réelles. In: Cahiers de l'Observatoire des pratiques linguistiques, 1, Orléans, (Délégation générale à la langue française et aux langues de France – Presses Universitaires d'Orléans).
- Biedermann-Pasques, L., & Jejcic, F. (2005): "Orthographe". In: Encyclopaedia Universalis, site abonnés internet: <http://www.universalis.fr>.
- Catach N., Gruaz, C. & Duprez, D. (1980, 2e éd. 1984, 3e éd. 1995): L'orthographe française, traité théorique et pratique. Paris (Nathan).
- Dictionnaire de l'Académie française (9e édition en cours), t. 1, A-Enzyme, 1992. Paris (Imprimerie Nationale); t. 2, Écène-Mappemonde, 2000, Paris (Fayard, Imprimerie Nationale); t. 3, Maquereau-Quotité, 2011 Paris (Fayard, Imprimerie Nationale); la suite du Dictionnaire paraît en fascicules, dernier paru en 2011, de *Quadru-* à *Raidisseur*, Paris Imprimerie Nationale, Documents Administratifs du Journal Officiel de la République Française.
- Fondet, C. (1990): Vaincre l'illettrisme. Paris (Édition Science et service Quart Monde).
- Martinet, A. & Walter, H. (1973): Dictionnaire de la prononciation française dans son usage réel. Paris (France-Expansion).
- Martinet, A., Martinet, J. et al. (1983): Vers l'écrit avec Alfonic. Paris (Hachette), rééd. Paris (RAPHAEL, 2006).
- Taulelle, D., Rey-Debove, J., Boumendil-Lucot, A., Lechheb, S. et al. (1989): Le Robert Oral-Écrit. Paris (Dictionnaires Le Robert).

Sites internet

<<http://www.inforef.be/projets/jeparledoncjecris/index.html>>, accès libre au didacticiel "Je parle donc j'écris".

<<http://www.vjf.cnrs.fr/orthographe>>, accès aux documents officiels concernant les rectifications de l'orthographe française et à la bibliographie thématique.

L'enseignement et l'apprentissage de la conjugaison en FLE: comment réduire les difficultés engendrées par l'orthographe?

Christian SURCOUF

École de français langue étrangère (EFLE), Université de Lausanne,
Faculté des lettres UNIL

French verb morphology has always been a major challenge for learners as well as teachers of French as a foreign language. Learning difficulties arise not only from the inherent complexity of the conjugation system itself, but mostly from the traditional description found in specialized books, grammars, etc. French spelling alone tends to complexify the actual oral verb morphology by more than 60%, thus hindering efficient learning. Following Dubois (1967), Csécsy (1968), Pouradier Duteil (1997), etc., I suggest an alternative approach, *exclusively* based on phonetic transcription, and starting with plural forms instead of singular ones (Mayer 1969). For more than 500 verbs of the 2nd and 3rd groups, this strategy allows learners to first memorize the present tense plural form e.g. /illiz/ (*ils lisent*, "they read") and take the stem's final consonant away to get the singular /illi/ (*il lit*, "he reads").

Introduction

Il suffit de parcourir la Toile pour constater que nombre de scripteurs francophones ont des difficultés en orthographe. Qu'un tel problème se retrouve chez les apprenants de français langue étrangère (FLE) n'a rien d'étonnant. La problématique est cependant différente. Alors que le scripteur francophone natif a déjà une excellente maîtrise de la langue avant d'appréhender le code écrit, il en est tout autrement des apprenants de FLE, pour qui, à la difficulté de l'apprentissage de la langue se greffe celle de l'acquisition de l'orthographe, notamment de sa lecture. D'un point de vue pédagogique, il paraît donc important de proposer des stratégies susceptibles d'atténuer cette double difficulté.

En nous concentrant sur l'apprentissage de la conjugaison, on examinera dans un premier temps les problèmes engendrés par la classification traditionnelle, fondée à la fois sur la nature de l'infinitif du verbe et les spécificités orthographiques de sa conjugaison. Une fois mises en évidence les diverses difficultés engendrées par une telle classification, on examinera l'intérêt pédagogique des propositions linguistiques basées sur la morphologie orale (dans la lignée de Csécsy, 1968; Dubois, 1967; Pouradier Duteil, 1997), en préconisant le recours à la transcription

phonétique, et une redistribution de l'ordre des personnes du paradigme, non plus de *je* à *ils*, comme le veut la tradition, mais de *vous* à *je*¹.

1. Les facteurs de complexité

Bien qu'au sein des langues romanes, la conjugaison du français s'avère une des plus simples, elle n'en reste pas moins complexe. Si dans la lignée du *français fondamental*, on ne retient que 339 verbes (Gougenheim *et al.*, 1964: 206) et qu'on exclut des temps simples le Passé Simple² et l'Imparfait du subjonctif, il reste néanmoins plus de 10 000 formes simples potentielles à acquérir. Si par ailleurs, "un mot sur quatre ou cinq est un verbe en français" (Catach, 1995: 227), la conjugaison s'impose comme incontournable, et constitue l'un des défis majeurs dans l'enseignement/apprentissage du FLE. Toute réflexion pédagogique sur la question doit par conséquent s'efforcer de repérer les facteurs de *complexification* et, si possible, proposer des solutions pour les atténuer ou, dans le meilleur des cas, les éliminer.

Il est important de différencier *complexité* de *complexification*. La complexité peut être inhérente à l'objet même, et/ou résulter de la présentation qui en est faite (i.e. d'une complexification). Par exemple, la morphologie verbale du français s'avère plus *complexe* que celle de l'anglais (le Présent anglais connaît deux formes du verbe *drink* /drɪŋk/ /drɪŋks/ contre quatre en français /bwa/, /byvõ/, /byve/, /bwav/)³. Ce type de complexité est irréductible. L'apprenant est contraint d'acquérir l'ensemble du paradigme en usage, quelle qu'en soit le degré de complexité. À cette complexité inhérente au fonctionnement même de la langue, s'ajoute celle de toute démarche métalinguistique de description, qu'elle soit scientifique ou pédagogique. Comme on le démontrera quantitativement, ce second type de complexité issu des stratégies de présentation de la conjugaison peut faire l'objet d'améliorations si, nous allons le voir, on écarte – au moins provisoirement – l'orthographe, et l'infinitif comme critères de classification.

¹ En ce sens notre proposition se démarque des précédentes (cf. note 23).

² Les temps grammaticaux, ou " tiroirs verbaux ", sont écrits avec une initiale majuscule, ou abrégés ainsi: Pr (Subj): Présent (du subjonctif), Imp: Imparfait, PS: Passé simple, Cond (Pr): Conditionnel Présent.

³ Parmi les facteurs de complexité d'une langue recensés par Nichols (2009: 111-112) apparaît le nombre de formes de la conjugaison (pour une réflexion critique sur la notion de "complexité" dans les langues, voir Deutscher, 2009).

2. La complexité due à l'orthographe

2.1. Le nombre de formes écrites et orales avant décomposition

Préoccupons-nous tout d'abord du nombre de formes écrites et orales à l'aide de deux exemples concrets, l'un "régulier" (à ce propos voir les réserves de Pouradier Duteil, 1997: 74), *ôter*, l'autre "irrégulier", *être*, conjugués à tous les temps simples usuels⁴. Pour l'instant, notre examen ne concerne que les formes écrites et orales globales, c'est-à-dire, en partie, ce à quoi l'apprenant se voit confronté dès qu'il ouvre un ouvrage de référence. En théorie, six personnes déclinées en six temps simples donneraient (6x6) 36 formes, mais qu'en est-il exactement (à ce propos, voir également les réflexions de Besse & Porquier, 1984: 51)?

	Pr	Imp	PS	Futur	Cond Pr	Pr Subj
1	ôte /ot/	ôtais ⁵ /otɛ/	ôtai /otɛ/	ôterai /otɛʁ/ ⁶	ôterais /otɛʁɛ/	ôte /ot/
2	ôtes /ot/	ôttais /otɛ/	ôtas /ota/	ôteras /otɛʁa/	ôterais /otɛʁɛ/	ôtes /ot/
3	ôte /ot/	ôtait /otɛ/	ôta /ota/	ôtera /otɛʁa/	ôterait /otɛʁɛ/	ôte /ot/
4	ôtons /otõ/	ôtions /otjõ/	ôtâmes /otam/	ôterons /otɛʁõ/	ôterions /otɛʁjõ/	ôtions /otjõ/
5	ôtez /otɛ/	ôtiez /otjɛ/	ôtâtes /otat/	ôterez /otɛʁɛ/	ôteriez /otɛʁjɛ/	ôtiez /otjɛ/
6	ôtent /ot/	ôtaient /otɛ/	ôtèrent /otɛʁ/	ôteront /otɛʁõ/	ôteraient /otɛʁɛ/	ôtent /ot/

écrit	5 formes	5 formes	6 formes	6 formes	5 formes	5 formes
oral	3 formes	3 formes	5 formes	3 formes	3 formes	3 formes

nombre total de formes écrites différentes:	27 (sur 36)
nombre total de formes orales différentes:	14 (sur 36)

Tableau 1 – Les 36 conjugaisons des temps simples du verbe ôter (sans l'Imparfait du subjonctif)

⁴ En accord avec Blanche-Benveniste (2010: 64), le Subjonctif Imparfait est considéré comme désuet, et est exclu du tableau. Selon *Lexique 3.71* (corpus comportant plus de 50 millions de mots selon New *et al.*, 2007: 663), la forme la plus courante de l'Imparfait du subjonctif est *fussions* (dans le corpus écrit) et n'apparaît qu'une fois toutes les 10⁴ occurrences de mots. L'Impératif, identique au Présent pour la quasi-totalité des verbes, est ici ignoré.

⁵ Même si une telle différence subsiste, entre autres, dans certaines régions de Suisse romande (Detey *et al.*, 2010: 223), comme le remarque Léon (2007: 89): "Les paires minimales telles que *irai* /iʁɛ/ ≠ *irais* /iʁɛ/ tendent à disparaître soit au profit de [e] en province, soit au profit de [ɛ] à Paris". Aussi neutralisera-t-on la différence à l'aide de l'archiphonème /ɛ/ (cf. également Brissaud *et al.*, 2006; Éluerd, 2009: 23).

⁶ En "français standard", le /ə/, selon les contextes, peut être facultatif (comme ici), obligatoire /ʒɑ̃tɛʁɛ/ (*j'entrera*), impossible /ʒlɛʁɛ/ (*je louera*) (cf. Gaatone, 2001: 216).

À l'opposé de cette régularité, quelle est la conjugaison d'*être*, à la fois verbe le plus irrégulier et le plus fréquent de la langue française (Le Goffic, 1997: 13)? Diffère-t-elle radicalement à cet égard?

	Pr	Imp	PS	Futur	Cond Pr	Pr Subj
1	suis /sɥi/	étais /etE/	fus /fy/	serai /sɛE/	serais /sɛE/	sois /swa/
2	es /E/	étais /etE/	fus /fy/	seras /sɛa/	serais /sɛE/	sois /swa/
3	est /E/	était /etE/	fut /fy/	sera /sɛa/	serait /sɛE/	soit /swa/
4	sommes /sɔm/	étions /etjō/	fûmes /fym/	serons /sɛō/	serions /sɛkjō/	soyons /swajō/
5	êtes /ɛt/	étiez /etje/	fûtes /fyt/	serez /sɛE/	seriez /sɛkjE/	soyez /swaje/
6	sont /sō/	étaient /etE/	furent /fyɛ/	seront /sɛō/	seraient /sɛE/	soient /swa/

écrit	6 formes	5 formes	5 formes	6 formes	5 formes	5 formes
oral	5 formes	3 formes	4 formes	3 formes	3 formes	3 formes

nombre total de formes écrites différentes:	32 (sur 36)
nombre total de formes orales différentes:	20 (sur 36)

Tableau 2 – Les 36 conjugaisons des temps simples du verbe être (sans l'Imparfait du subjonctif)

Afin de mieux percevoir la différence entre les présentations écrite et orale et la complexification que l'une ou l'autre peut engendrer lors de l'apprentissage, synthétisons les résultats des deux tableaux en comparant maintenant les écarts entre le nombre de formes écrites et orales pour chacun des deux verbes. Dans tous les cas, le passage de l'oral à l'écrit implique un accroissement de 20% à 100% du nombre de formes dans les deux verbes choisis.

Le verbe <i>ôter</i>	Pr	Imp	PS	Fut	Cond	Pr Sbj
nb de formes écrites	5	5	6	6	5	5
nb de formes orales	3	3	5	3	3	3
<i>écrit</i> → oral	-40%	-40%	-17%	-50%	-40%	-40%
oral → <i>écrit</i>	+67%	+67%	+20%	+100%	+67%	+67%

nb total de formes écrites différentes:	27	moyenne par tiroir verbal	5,33
nb total de formes orales différentes:	14	moyenne par tiroir verbal	3,33
<i>écrit</i> → oral	-48%		
oral → <i>écrit</i>	+93%		

Tableau 3 – Variations du nombre de formes dans le passage de l'écrit à l'oral ou de l'oral à l'écrit (*ôter*)

Le verbe <i>être</i>	Pr	Imp	PS ⁷	Fut	Cond	Pr Sbj
nb de formes écrites	6	5	5	6	5	5
nb de formes orales	5	3	4	3	3	3
<i>écrit</i> → oral	-17%	-40%	-20%	-50%	-40%	-40%
oral → <i>écrit</i>	+20%	+67%	+25%	+100%	+67%	+67%

nb total de formes écrites différentes:	32	moyenne par tiroir verbal	5,33
nb total de formes orales différentes:	20	moyenne par tiroir verbal	3,5
<i>écrit</i> → oral	-38%		
oral → <i>écrit</i>	+60%		

Tableau 4 – Variations du nombre de formes dans le passage de l'écrit à l'oral ou de l'oral à l'écrit (*être*)

Si le recours au pourcentage peut paraître abusif pour représenter des variations en deçà de 6, la disparité entre écrit et oral exposée ici est extrapolable – avec quelques ajustements selon les verbes – à *toute* la conjugaison du français. En d'autres termes, une disparité potentielle comprise entre 60% (*être*) et 93% (*ôter*) entre formes orales et écrites complexifie la morphologie orale d'autant sur des centaines de verbes. En somme, l'apprenant consultant un ouvrage de référence ordinaire percevra la conjugaison du français comme au moins 60% plus complexe qu'elle ne l'est en réalité à l'oral.

Par ailleurs, à ses débuts, quel que soit le système scriptural de sa langue source, il est peu probable que l'apprenant saisisse toute la complexité de l'enchevêtrement des quatre mécanismes régissant l'orthographe du français (phonogrammique, morphogrammique, distinctif, et étymologique, cf. Catach, 1995: 27s; 2004: 53s; Riegel *et al.*, 2009: 121s). Si, à l'instar de tout système alphabétique, l'orthographe du français repose en grande partie sur un principe phonogrammique (80 à 85% selon Catach, 1995: 27)⁸, parmi les trois restants, le principe morphogrammique rend délicate la compréhension du fonctionnement de la conjugaison orthographique, et son extrapolation à l'oral. En effet, si certains morphogrammes ont un pendant oral, à l'instar du digramme *ai* se lisant /E/ dans l'Imparfait // *dansait* (Riegel *et al.*, 2009: 127), d'autres n'offrent aucune contrepartie orale. Dans /ildās(t)isi/ *Ils dansent ici*, alors que la liaison facultative

⁷ On conserve ici le nombre théorique. Les formes d'adresse /tyota/ (*tu ôtas*), /tyfy/ (*tu fus*) et /vuzotat/ (*vous ôtâtes*), /vufyt/ (*vous fûtes*) étant inusitées, le nombre effectif de formes orales (surtout lues) se réduirait en définitive à 4 pour *ôter* et 3 pour *être*.

⁸ Catach (1995: 36) recense 133 phonogrammes susceptibles d'être réduits à 45 dans un "code minimal de transcription du français, nécessaire à l'expression écrite d'un scripteur débutant" (réduits à 33 archigraphèmes chez Riegel *et al.*, 2009: 122). La problématique est toutefois différente, puisqu'en FLM, les apprentis scripteurs connaissent déjà la plupart des formes orales des mots à écrire et à lire. Tel n'est pas le cas en FLE, où les irrégularités parasitent l'association entre formes orale et écrite, et rendent délicate la recherche de toute logique de fonctionnement (cf. par ex. *lien* /ljɛ̃/, vs *lient* /li/, *tien* /tjɛ̃/ et *tient* /tjɛ̃/, etc.).

parvient occasionnellement à justifier la présence du *t* final, le *n* demeure quant à lui inexpliqué.

Dans la lecture des formes orthographiques de la conjugaison, les débutants en quête de cohérence tendent à généraliser le principe phonogrammique en attribuant un phonème à chaque graphème. Une telle attitude peut en partie résulter d'une projection des mécanismes du système orthographique de la langue source (s'il existe). On peut par ailleurs supposer que cette tendance est d'autant plus forte que ce système orthographique source s'approche d'une biunivocité graphème/phonème⁹. Par exemple, les apprenants hispanophones (mais pas seulement) tendent à lire erronément (*tu*) *vas* /vas/, *as* /as/ (formes qui, en l'occurrence, font écho à celles de l'espagnol), (*ils*) *font* /fɔ̃t/, *sont* /sɔ̃t/¹⁰, etc.

Conséquence directe de la présence de ces "lettres muettes", plusieurs formes écrites différentes peuvent transcrire une forme orale unique: /kuʁ/ s'écrit *cours*, *court*, *courent* au Présent de l'Indicatif et *coures*, *coure*, et aussi *courent* au Subjonctif (sur la difficulté des apprenants de FLE à dissocier les formes indicatives et subjonctives, cf. Besse & Porquier, 1984: 51). Parmi les verbes "irréguliers", d'autres singularités orthographiques, plus ponctuelles, contribuent également à opacifier le fonctionnement de la conjugaison. Constatant la présence du *d* dans *Il défend*, *il fond*, etc. l'apprenant pourrait supposer qu'il sert d'indicateur, à la fois écrit et oral, du pluriel *ils défendent* /ildefãd/, *ils fondent* /ilfõd/, etc. Des contre-exemples existent cependant: *il prend*, *il coud*, etc. ne donnent pas *ils *prendent*, *ils *coudent* (du verbe *couder*, pas *coudre*). Une seconde hypothèse consisterait à justifier la présence de ce *d* à l'aide de l'infinitif: *défendre*, *fondre*, *coudre*, *prendre*, etc. Mais là encore les verbes en *-indre* invalident un tel raisonnement: *il craint*, *il joint*, *il peint* (cf. les remarques de Catach, 1995: 234; Thimonnier, 1967: 257-258). De même, alors que la logique conduira l'apprenant à une lecture correcte de *vit-il* /vitol/, *fait-il* /fetil/, elle le mènera à lire incorrectement *prend-il* */pʁãdil/, *s'assied-il* */sasjedil/. Si l'orthographe pose des problèmes au francophone natif – même expert – en production (Jaffré, 2005: 354-355), elle ne peut néanmoins jamais le conduire à inférer que le pluriel de *il coud* serait *ils coudent*. Tel n'est évidemment pas le cas de l'apprenant de FLE, qui

⁹ Katz & Frost (1992: 69) citent par ex. l'orthographe du serbo-croate: "Each letter represents only one phoneme and each phoneme is represented by only one letter". Le hongrois, l'espagnol, l'italien, le finnois, etc. présentent des fonctionnements proches de la biunivocité. Si l'orthographe française s'illustre comme l'une des plus complexes (Jaffré, 2005: 354-355), alors la probabilité d'une projection simplificatrice chez les apprenants est élevée.

¹⁰ La langue espagnole dicterait ici /son/, sans /t/.

s'évertuera à chercher des régularités sur la base des informations orthographiques à sa disposition.

Bien entendu, comme le montrent les recherches sur l'orthographe et sa didactique (par ex. Catach, 1995; Catach, 2004; Thimonnier, 1967), la plupart de ces graphies peuvent faire l'objet de justifications. Toutefois, si en français langue maternelle (FLM), il est possible de "tabler sur la connaissance, relativement approfondie, qu'ils [=les élèves francophones natifs] ont de la langue parlée" (Thimonnier, 1967: 255), en FLE, les débutants se retrouvent quant à eux confrontés à l'apprentissage simultané de la langue orale et écrite, sans que la première puisse prévenir les erreurs de lecture de la seconde. Au contraire, les difficultés d'appariement entre formes orales et écrites rendront toujours plus délicate la quête de compréhension du fonctionnement du système morphologique.

2.2. *La complexité des modèles proposés dans les ouvrages de référence*

À de rares exceptions près, les manuels, les grammaires, les dictionnaires – bilingues ou monolingues –, reproduisent la classification traditionnelle exposée dans les ouvrages spécialisés, dont nous allons examiner quatre exemples ici: le *Bescherelle* (Arrivé, 2006), le *Larousse* (Le Lay, 1995), le *Littré* (Éluerd, 2009) ou le *Robert & Nathan* (Carelli *et al.*, 1995). Sans qu'elle soit toujours ouvertement déclarée, la finalité de tels manuels est avant tout orthographique. Ainsi le *Robert & Nathan* précise-t-il que "les difficultés spécifiques [sont] mises en évidence par des jeux de couleurs dans chaque tableau" (1995: 3), et surligne par exemple le *e* dans "nous mangeons", signalant manifestement non pas une "difficulté spécifique" de la conjugaison française mais de son orthographe (tous les francophones savent conjuguer *manger*).

Rappelons une évidence. Dans le fonctionnement de toute langue vivante, l'oral est premier, et l'écrit, s'il existe, en est la transposition plus ou moins cohérente¹¹. En d'autres termes, le fonctionnement systémique d'une langue provient de son usage oral. La morphologie verbale n'échappe pas à cette loi¹². Dès lors, en apprentissage, la conjugaison orale devrait nécessairement primer. Si les irrégularités orthographiques peuvent intéresser les apprenants de FLE, leur effort devra cependant porter au

¹¹ En d'autres termes, toute langue vivante est acquise par le bébé de manière orale. L'écrit, s'il existe, ne précède jamais l'oral. C'est en ce sens que l'oral est premier.

¹² Citons l'exception de *argüer*, dont l'orthographe *arguer* avant 1990, est probablement à l'origine de conjugaisons telles que /ʒaʁg/, /nuzaʁgõ/, rimant avec celle de *narguer*, là où l'on attendrait une conjugaison de type *dénuer*, conformément aux paires *dénuer-dénuelement*, *argüer-argument*.

premier chef sur la compréhension des mécanismes élémentaires de la morphologie verbale *orale*, dont la restitution orthographique, nous l'avons vu en partie, tend à occulter les régularités.

Examinons d'autres écueils de la présentation traditionnelle, et penchons-nous sur le "1^{er} groupe", que les quatre manuels retenus subdivisent en un nombre variable de sous-groupes (N), chacun représenté par un verbe-modèle:

	N	verbe modèle ¹³
<i>Bescherelle</i>	13	aimer, placer, manger, peser, céder, jeter, modeler, créer, assiéger, apprécier, payer, broyer, envoyer
<i>Larousse</i>	22	aimer, créer, étudier, distinguer, manger, placer, acquiescer, céder, protéger, rapiécer, appeler, interpellier, geler, dépecer, jeter, acheter, payer (1), payer (2), employer, essuyer, envoyer, arguer
<i>Littre</i> ¹⁴	31	chanter, briller, signer, créer, crier, étudier, jouer, saluer, payer (1), payer (2), grasseyer, envoyer, nettoyer, essuyer, aider, pleurer, commencer, acquiescer, manger, déneiger, arguer, dépecer, geler, appeler, interpellier, acheter, jeter, lever, rapiécer, protéger, céder
<i>Robert & Nathan</i>	34	parler, commencer, acquiescer, marquer, manger, distinguer, arguer, payer, grasseyer, employer, envoyer, ennuyer, prier, signer, briller, pallier, travailler, réveiller, mouiller, semer, dépecer, acheter, jeter, geler, appeler, interpellier, créer, accéder, abréger, régner, déléguer, disséquer, remuer, louer

Tableau 5 – Les subdivisions du "1^{er} groupe" dans quatre ouvrages de référence

On remarquera en premier lieu la divergence considérable dans le nombre de sous-groupes selon les ouvrages, allant presque du simple au triple. Quel qu'en soit le degré de minutie, ces subdivisions tendront inévitablement, en FLE, à entraver le repérage des régularités de la conjugaison orale. En effet, certains verbes, à l'instar d'*acquiescer*, *briller*, *commencer*, *naviguer*, *bouger*, *marquer*, etc. présentés comme différents – en vertu de particularités orthographiques – obéissent à un modèle de conjugaison orale unique et parfaitement régulier de type *parler*. En voici l'illustration pour le Présent:

ʒ(ə), ty, il(z)	paʁl, akʁɛs, bʁij, komãs, navig, buʒ, maʁk	–
nu(z)		õ
vu(z)		e

Tableau 6 – Le Présent régulier de verbes relevant de sous-groupes orthographiques différents

¹³ On conserve ici l'orthographe des auteurs, qui justifie à elle seule l'existence des sous-groupes pour *arguer*, et *interpeller* (*argüer* et *interpeler* après 1990).

¹⁴ Le *Littre* propose une transcription phonétique des bases orales, et fusionne des approches orthographiques et orales. Par ex. la conjugaison de *jouer* se différencie de celle de *chanter* en raison de "deux radicaux à l'oral [ʒu-], [ʒw-]", *pleurer*, parce qu'il connaît une alternance [œ/ø] (Éluerd, 2009: 106 & 115), etc.

Si la répartition en sous-groupes sur des fondements uniquement orthographiques établit des discriminations sans pertinence pour l'oral, ce n'est cependant pas toujours le cas. En effet, certains sous-groupes sont le reflet de véritables différences dans la morphologie orale. Tel est le cas par exemple des verbes modèles *céder*, *acheter*, *broyer*.

ɜ(ə), ty, il(z)	sɛd, aʃɛt, bʁwa	–
nu(z)	sed, aʃ(ə)t, bʁwaj	ð
vu(z)		e

Tableau 7 – Le Présent à deux bases¹⁵ de verbes relevant de sous-groupes orthographiques différents

Comme l'illustre ce bref examen du 1^{er} groupe, pour l'apprenant de FLE en quête de compréhension de la morphologie orale, la consultation des ouvrages spécialisés de conjugaison (ou de dictionnaires, manuels, grammaires) pose au moins deux problèmes. La répartition en sous-groupes orthographiquement motivés ne permet pas d'inférer le fonctionnement de la morphologie orale, puisque deux sous-groupes peuvent *ou non* avoir une conjugaison orale similaire. Par ailleurs, cette classification donne une vision beaucoup plus complexe de la réalité morphologique. Dans le cas extrême du *Robert & Nathan*, les verbes à conjugaison orale unique (*aimer*, *briller*, *marquer*, etc.) se voient répartis en 19 sous-groupes pour des raisons orthographiques (18 pour *Littré*, 9 pour *Larousse*, 4 pour *Bescherelle*). Le tableau suivant synthétise les cas de figure des quatre ouvrages pour chaque type de verbe modèle du 1^{er} groupe (à 1 base du type *marquer* /maʁk/; à 2 bases du type *lever* /lɛv/-/lɛv-/, *céder* /sɛd/-/sed-/, *broyer* /bʁwa/-/bʁwaj-/, et à 3 bases du type *envoyer* /ãvwa/-/ãvwaj-/-/ãveʁ-/).

nb de bases dans le verbe modèle:	1	2			3
		ə/ɛ	ɛ/e	-/j	
nb de sous-groupes correspondants dans le:					
<i>Bescherelle</i>	4	3	3	2	1
<i>Larousse</i>	10	6	3	2	1
<i>Littré</i>	18	7	3	2	1
<i>Robert & Nathan</i>	19	7	5	2	1

Tableau 8 – Correspondances entre sous-groupes de conjugaisons orale et écrite dans 4 ouvrages

Dans le pire des cas, l'apprenant se retrouve confronté à une conjugaison orthographique paraissant dix-neuf fois plus complexe qu'elle ne l'est en réalité à l'oral. Il est alors probable que l'apprenant soucieux de comprendre les régularités de la morphologie verbale se sente découragé

¹⁵ Pour ce terme, cf. §4.

face à une telle complexité, et à son manque de cohérence par rapport à l'oral.

En définitive, comme le montre ce bref examen, il semble difficile en début d'apprentissage de proposer une vision systémique claire, simple, et immédiatement accessible de la présentation orthographique de la conjugaison.

3. La complexité de la classification traditionnelle en dehors de l'orthographe

Si la dimension orthographique entrave considérablement toute recherche de systémicité, là ne réside pas le seul inconvénient de la classification traditionnelle. On l'a vu, un verbe en *-er*, relevant par définition du 1^{er} groupe (à l'exception d'*aller*), n'obéit pas nécessairement à un modèle unique de conjugaison, et peut correspondre en fait à cinq types différents (dont un ne concerne que *(r)envoyer*). Qu'en est-il maintenant des deux autres groupes? Le 2^e est le plus homogène de la conjugaison française, il se construit sur deux bases orales (par ex. /finis/-/fini/), et ne comporte en français continental qu'une seule exception: *hair*¹⁶. En revanche, le *Robert & Nathan* décourage d'emblée toute quête de logique pour le 3^e groupe, puisque, selon les auteurs, "il n'existe pas de modèle de conjugaison et donc de règle de formation pour le 3^e groupe" (Carelli *et al.*, 1995: 13). Si, on le verra, cette affirmation est à nuancer, il n'en demeure pas moins que la classification traditionnelle inspirée du latin (Dubois, 1967: 56; Germain & Séguin, 1998: 11) et articulée sur l'infinitif pose plusieurs problèmes pédagogiques en FLE.

En premier lieu, en dépit des apparences, le choix de l'infinitif comme indice de classification ne permet pas de prédire avec certitude l'appartenance à un groupe¹⁷. C'est notamment le cas des verbes en *-ir*, susceptibles de relever du 2^e ou du 3^e groupe. Ainsi sur les 400 verbes en *-ir* du *Petit Robert* (2011), 313 (78%) appartiennent au 2^e groupe (type *finir*), 87 (22%) au 3^e, répartis en dix types de conjugaisons différents, avec comme verbes modèles: *acquérir, assaillir, courir, couvrir, cueillir, fuir, mourir, sentir, tenir, vêtir* (cf. Kilani-Schoch & Dressler, 2005: 166-169). Cette opacité du classement se voit aggravée par la présence de 61 verbes du 3^e groupe avec une rime similaire en /iʁ/, mais une orthographe en *-ire*,

¹⁶ Le paradigme est régulier pour de nombreux canadiens francophones, avec deux bases /ai/-/ais/.

¹⁷ Pouradier Duteil (1997: 8) préconise l'usage de la 3^e personne du pluriel du Présent à la place de l'infinitif (cf. cependant la défense de l'infinitif par Kilani-Schoch & Dressler, 2005: 151-152). Gaatone (2001: 217) rappelle qu'il n'est pas toujours possible "de prévoir le nombre et la forme des bases à partir de l'une d'entre elles" (pour une alternative, cf. également les réflexions de Bonami & Boyé, 2007).

dépendant de six types de conjugaison avec pour verbes modèles *cuire*, *écrire*, *luire*, *dire*, *maudire* et enfin *rire*, dont la conjugaison orale est identique à celle de *fuir* (verbe en *-ir*). Récapitulons. Sur l'ensemble des 6628 verbes du *Petit Robert* (2011), on obtient la répartition suivante (cf. également Séguin, 1989: 125):

	1 ^{er} Gr	2 ^e Gr	3 ^e Gr	Total
nb de verbes	5922	313	393	6628
%	89%	5%	6%	100%

Tableau 9 – Répartition des verbes dans les groupes dans le *Petit Robert* (2011)

Si l'on exclut *hair*, seul le 2^e groupe est homogène. Cependant, en raison d'une porosité entre les 2^e et 3^e groupes, l'apprenant de FLE face à un verbe en *-ir* court le risque de faire une prédiction fausse dans un cas sur cinq, et par exemple produire *nous *offrissons, je *courirai* sur le modèle de *finir, nous finissons, je finirai*. Une difficulté analogue, mais de moindre ampleur, existe avec 500 verbes (8%) du 1^{er} groupe, qui n'obéissent pas à la conjugaison à base unique des 92% restants de type *chanter* (cf. Tableau 8). Le problème essentiel concerne donc le 3^e groupe, qui ne comprend que 6% des verbes, dont beaucoup sont très courants. Ainsi, en excluant *être* et *avoir*, parmi les 100 verbes les plus fréquents de la base *Lexique 3.71* (New et al., 2007) des sous-titres de films, 49 appartiennent au 3^e groupe, 2 au 2^e, et 49 au 1^{er}. En somme, les verbes du groupe dont la conjugaison est à priori la moins prédictible et la plus difficile à apprendre comptent pour la moitié des occurrences.

Si, dans sa quête de compréhension des régularités orales de la conjugaison, l'apprenant recherchait une forme de systémicité au sein de la classification traditionnelle – de loin la plus courante dans l'enseignement/apprentissage du FLE et du FLM –, il pourrait émettre des hypothèses selon plusieurs axes, et supposer que:

- (a) La répartition en trois groupes se voit justifiée par un fonctionnement homogène au sein de chaque groupe;
- (b) L'infinitif est un indicateur univoque du groupe et du type de conjugaison qui constitue ce groupe;
- (c) Chaque groupe a un nombre fixe de bases orales.

Or, on l'a vu, aucune de ces hypothèses n'est vérifiée. Ajoutons que, contre l'hypothèse (c), des verbes tels que *tuer*, et *conclure* se trouvent respectivement dans le 1^{er} et le 3^e groupe en dépit de conjugaisons très similaires. En effet, si l'on exclut le Participe passé, on a dans les deux cas une base unique pour construire l'ensemble du paradigme: /ty/ et /kōkly/. Pour l'apprenant débutant (ou avancé), ce genre de rapprochement, bien que souhaitable, s'avère d'autant plus difficile qu'il va à l'encontre de

toutes les présentations traditionnelles disponibles. Par ailleurs, pédagogiquement, on ne peut s'attendre à ce que l'apprenant interroge la pertinence des regroupements proposés dans les "ouvrages de référence", dont le statut d'autorité repose justement sur la confiance que leur confèrent les usagers. En définitive, toute attitude de défiance contreviendrait à la dynamique même de consultation de tels ouvrages.

Maintenant que les problèmes essentiels¹⁸ sont identifiés, que proposer comme alternative susceptible de simplifier ou, du moins, d'éviter de complexifier l'apprentissage de la conjugaison?

4. Présentation linguistique succincte de la conjugaison

Ne sera ici fournie qu'une présentation linguistique succincte concernant la morphologie verbale du français (des développements plus ou moins conséquents sont proposés par Blanche-Benveniste, 2010: 63-78; Bonami & Boyé, 2007; Boyé, 2000; Csécsy, 1968; Dubois, 1967: 56-79; Gaatone, 2001; Germain & Séguin, 1998: 70-81; Kilani-Schoch & Dressler, 2005: 117-210; Le Goffic, 1997; Mahmoudian, 1976: 310-318; Marty, 1971: 105-117; Pinchon & Couté, 1981; Pouradier Duteil, 1997; Séguin, 1989: 124-137; Touratier, 1996: 15-58; Wagner & Pinchon, 1991: 262-269). Même si d'un point de vue linguistique, "après un siècle d'analyse ou presque, aucun consensus ne semble près de se dégager" (Morin, 2008: 136), les regroupements s'avèrent pédagogiquement plus satisfaisants que la classification traditionnelle, dans la mesure où ils s'effectuent sur la base de l'oral, et évitent les écueils orthographiques mentionnés plus haut.

L'intérêt de la décomposition des formes orales repose sur un ensemble de mécanismes permettant d'isoler d'une part un nombre restreint de *bases* (ici jusqu'à cinq pour les non-supplétifs¹⁹), d'autre part les *désinences*. En FLE, si, pour les niveaux débutants, on exclut le Passé simple (cf. Gougenheim *et al.*, 1964: 218) et l'Imparfait du Subjonctif de la présentation, on obtient cinq tiroirs verbaux simples usuels, auquel il faudrait adjoindre le Participe passé pour autoriser la construction des temps composés²⁰. La conjugaison du verbe *ôter* donnerait le tableau suivant:

¹⁸ Faute de place, nous n'avons pas évoqué les problèmes de la décomposition des formes orthographiques, en radical et désinences.

¹⁹ C'est-à-dire, très schématiquement (pour une discussion de cette problématique, qui dépend notamment du cadre théorique adopté, cf. Bonami & Boyé, 2003), des verbes tels que *avoir*, *être*, *aller*, qui sont à la fois indécomposables selon ces principes, et comportent de surcroît des bases sans relation morphologique immédiate entre elles, telles que par ex. pour *aller* /va/, /alō/, /iʁa/, /aj/, etc. (cf. Kilani-Schoch & Dressler, 2005: 72).

²⁰ L'Impératif est identique au Présent de l'indicatif, sauf pour *avoir*, *être*, *savoir*, *vouloir*, où il reprend la base du Présent du Subjonctif.

	Pr	Imp	Futur	Cond Pr	Pr S
1					
2		E		(ə)ʁ E	
3					
4	ot	ot	ot (ə)ʁ	ot (ə)ʁ	ot
5		j		əʁ j	j
6		E		(ə)ʁ E	E

Tableau 10 – La conjugaison orale du verbe *ôter* des cinq tiroirs verbaux simples usuels

En renouvelant la même opération sur une multitude de verbes, apparaît alors la possibilité de séparer d'une part une constante, ici /ot/, constituant la *base*, et de l'autre des *désinences* (à cet égard, cf. les réserves de l'approche par commutation formulées par Morin, 2007: 136). Si l'on isole les désinences pour l'ensemble des verbes du français (sauf pour *avoir*, *être*, *dire*, *faire*, *aller*), en tenant compte de la neutralisation de la différence /e/-/ɛ/ (cf. note 5), on aboutit à neuf désinences différentes²¹: /ō/, /E/, /jō/, /jE/, /(ə)ʁE/, /(ə)ʁa/, /(ə)ʁō/, /(ə)ʁjE/ et /(ə)ʁō/ (cf. également Le Goffic, 1997: 8; Marty, 1971: 107s) (pour une répartition sémantique des désinences en fonction des paramètres "personne", "nombre", "temps", et "modalité", cf. Gaatone, 2001: 217-218):

	Pr	Imp	Futur	Cond Pr	Pr S
1					
2		E			
3					
4	ō	j	(ə)ʁ	(ə)ʁ	j
5	E	E		j	E
6		E		E	

Tableau 11 – Grille désinentielle orale type des cinq tiroirs verbaux simples usuels

Ainsi devient-il possible de donner une présentation transversale de la conjugaison, en recourant à cette grille de désinences (ou "grille flexionnelle" chez Boyé, 2000: 37s). Limitons-nous à deux exemples:

²¹ On ne compte pas ici la "désinence zéro" (Le Goffic, 1997: 9) – comme dans /ilot-/ (*il ôte*) –, puisqu'elle est par définition inaudible, et ne présente guère de pertinence *pédagogique*. Certains phénomènes phonotactiques réguliers pourront être signalés dans le tableau individuel de chaque verbe. Par ex. à l'Imparfait /kōklyjō/ (*concluions*) s'inscrit parfaitement dans le gabarit, mais pas /ātɕijō/ (*entrions*), qui requiert un /i/ intercalaire, en raison d'une base avec une coda à double consonne /ātɕ/ (pour les autres "sons intercalaires", cf. Pouradier Duteil, 1997: 17-24).

	Pr	Imp	Futur	Cond Pr	Pr S
1					
2		E		(ə)ʁ E	
3	lav		lav (ə)ʁ		lav
4	õ	j õ	õ	əʁ j õ	j õ
5	E	E	E	E	E
6		E	õ	(ə)ʁ E	

Tableau 12 – Conjugaison orale du verbe *laver* aux cinq tiroirs verbaux simples usuels

	Pr	Imp	Futur	Cond Pr	Pr S
1					
2	pø	E		E	
3		puv	pu ʁ	pu ʁ	pqis
4	õ	j õ	õ	j õ	j õ
5	E	E	E	E	E
6	pœv	E	õ	E	

Tableau 13 – Conjugaison orale du verbe *pouvoir* aux cinq tiroirs verbaux simples usuels

En tant que tiroir verbal le plus utilisé (Maingueneau, 1999: 529; Riegel *et al.*, 2009), le plus complexe, et de surcroît fondamental dans la formation d'autres temps grammaticaux (Imparfait, Subjonctif Présent, et, selon les verbes, l'Impératif, le Futur et le Conditionnel Présent), le Présent occupe une position centrale et nécessairement prioritaire dans l'apprentissage du français. Pédagogiquement, c'est donc par lui qu'une telle approche orale de la conjugaison doit être mise en place. Si l'on exclut les verbes défectifs, et *être*, *avoir*, *aller*, *dire*, *faire*, chaque verbe s'inscrit dans un des cinq gabarits suivants de répartitions des bases, chacun illustré par un exemple:

Type I	Type II	Type III	Type IV	Type V
1 base <i>laver</i>	<i>jeter</i>	2 bases <i>peindre</i>	<i>finir</i>	3 bases <i>boire</i>
1				
2	3et	etɛ̃	fini	bwa
3				
4	õ	õ	õ	õ
5	3et	etɛ̃	finis	byv
6	3et			bwav

Tableau 14 – Les cinq gabarits de la conjugaison du Présent (hors *être*, *avoir*, *aller*, *dire*, *faire*)

Quantitativement, les 6577 verbes du *Petit Robert* (2011) retenus (des exemples figurent pour chaque type), se répartissent de la sorte (cf. également Séguin, 1989):

	Type I	Type II	Type III	Type IV	Type V
1 ^{er} Gr.	<i>laver, entrer, marcher</i>	<i>lever, céder, payer, noyer</i>			
total (% Gr)	5422 (92%)	500 (8%)	0	0	0
2 ^e Gr.				<i>finir, rougir, choisir</i>	
total (% Gr)	0	0	0	311 (100%)	0
3 ^e Gr.	<i>courir, offrir, conclure</i>	<i>quérir, mourir, croire</i>	<i>craindre, savoir, valoir</i>	<i>vivre, tendre, fondre</i>	<i>vouloir, tenir prendre</i>
total (% Gr)	32 (9%)	23 (7%)	37 (11%)	197 (58%)	55 (15%)
TOTAL	5454	523	37	508	55
	82,9%	8%	0,6%	7,7%	0,8%

Tableau 15 – La répartition des verbes du *Petit Robert* (2011) selon les cinq gabarits du Présent

Les types I, II et IV constituent environ 99% de l'ensemble de la conjugaison. Un verbe sur quatre se conjugue sur une base unique (I). Les 500 verbes de type II du 1^{er} groupe s'avèrent hautement prédictibles en fonction de leur rime à l'infinitif (cf. les répartitions du Tableau 8). À l'instar de tous ceux du 2^e groupe (sauf *hair*), la majorité des verbes du 3^e groupe (58%, 197 verbes) relèvent du gabarit IV, et se caractérisent par une base longue et une courte. À la suite de Séguin (1989: 129), on pourrait a priori considérer que "la marque du pluriel se fait par *adjonction* d'une consonne au radical nu" (nous soulignons), répertoriées ci-dessous:

CONSONNE	NB DE VERBES	EXEMPLE AU SINGULIER	EXEMPLE AU PLURIEL
/s/	328 (311+17)	/fini/ (<i>fini</i>), /nɛ/ (<i>nai</i>)	/finis/ (<i>finissent</i>), /nɛs/ (<i>naissent</i>)
/z/	53	/li/ (<i>li</i>)	/liz/ (<i>lisent</i>)
/d/	49	/tā/ (<i>tend</i>)	/tād/ (<i>tendent</i>)
/t/	46	/sā/ (<i>sent</i>)	/sāt/ (<i>sentent</i>)
/v/	22	/sɥi/ (<i>sui</i>)	/sɥiv/ (<i>suivent</i>)
/m/	3	/dɔʁ/ (<i>dort</i>)	/dɔʁm/ (<i>dorment</i>)
/p/	3	/ʁɔ/ (<i>rompt</i>)	/ʁɔp/ (<i>rompent</i>)
/k/	2	/vɛ̃/ (<i>vainc</i>)	/vɛ̃k/ (<i>vainquent</i>)
/j/	1	/bu/ (<i>bout</i>)	/buj/ (<i>bouillent</i>)
/l/	1	/mu/ (<i>moud</i>)	/mul/ (<i>moulent</i>)
total	508		

Tableau 16 – Répartition quantitative des dix consonnes²² des bases longues des verbes de type IV, issus du *Petit Robert* (2011)

En FLE, la *Grammaire des premiers temps* (Abry & Chalaron, 1997: 28-29) recourt à cette stratégie de présentation et explique que "la base

²² Contrairement à Séguin (1989: 129) et Kilani-Schoch & Dressler (2005: 147), on ne retiendra pas /ɲ/, qui, conjointement, met en œuvre une dénasalisation /pɛ̃l-/pɛɲ/ (*peint-peignent*) ou d'autres irrégularités /ʒwɛ̃l-/ʒwɛɲ/ (*joint-joignent*), et ne fonctionne donc pas comme /vɛ̃l-/vɛ̃k/, ou /tā-/tād/, etc.

longue=base courte + consonne". De nombreux exemples illustrent cette équation, à l'image de la présentation de *finir*, et des verbes s'inscrivant dans ce modèle caractérisé par l'ajout de "[s]":

FINIR			
je	fini	s	[fini]
tu	fini	s	
il	fini	t	
ils	finiss	ent	[finis]
vous	finiss	ez	
nous	finiss	ons	

- grandir, vieillir, rougir, applaudir
- paraître, disparaître, comparaître, naître, (re)connaître
- croître et ses composés
- maudire

Tableau 17 – Exemple de présentation de la conjugaison de *finir* (Abry & Chalaron, 1997: 28)

D'autres tableaux synthétiques sont présentés pour "+[z]", avec *taire* [tɛ]→[tɛz] comme modèle, "+[t]", avec *sortir* [sɔʁ]→[sɔʁt], etc. En dépit de son intérêt et de son aspect novateur en FLE, cette approche pourrait être améliorée. Imaginons un apprenant qui aurait intégré la règle selon laquelle à la base du singulier – première dans la présentation, donc plus saillante –, on ajoute /s/ comme dans /kʊʒi/, /fini/, /nɛ/, etc., alors, en présence de formes du singulier telles que /vi/ (*vit*), /li/ (*lit*), /plɛ/ (*plait*), présentant des rimes analogues, il pourrait faire l'hypothèse qu'il faut ajouter un /s/. Il aboutirait à des prévisions fausses */vis/, */lis/, */plɛs/. Pour éviter cet inconvénient, notre proposition consistera à inverser l'ordre des personnes en 5, 4, 6, 3, 2, 1²³. Ainsi, en présentant la base du pluriel en premier /liz/, /viv/, /plɛz/, etc. on permet à l'apprenant de trouver *avec certitude* la base du singulier en *soustrayant* la consonne finale. S'il a mémorisé que le verbe est à deux bases, de /liz/, /viv/, /plɛz/, etc. il obtiendra sans erreur /li/, /vi/, /plɛ/. Voici quelques présentations possibles du Présent, tenant compte de cette nouvelle disposition:

LAVÉ			COURIR			CONCLURE			JETER			NOYER		
vu	lav	e	vu	kʊʁ	e	vu	kōkly	e	vu	ʒɛt	e	vu	nwaʒ	e
nu		ō	nu		ō	nu		ō	nu		ō	nu		ō
il			il			il			il	ʒɛt		il	nwa	
il			il			il			il			il		
ty			ty			ty			ty			ty		
ʒə			ʒə			ʒə			ʒə			ʒə		

²³

Mayer (1969: 558-559) proposait déjà de mettre le pluriel en premier, mais dans l'ordre: *nous*, *vous*, *ils*. Toutefois, la forme avec *nous* disparaissant progressivement au profit de celle avec *on* (cf. Detey *et al.*, 2010: 98), il paraît préférable de présenter la forme avec *vous*, d'un usage beaucoup plus fréquent, en tête de paradigme. Dans l'apprentissage par cœur de la conjugaison – largement pratiqué –, cette forme, mémorisée en premier, sera plus immédiatement disponible pour l'interaction orale. De plus, pour les verbes en -er, (89% de la totalité), la forme avec *vous* est homophone de l'infinitif de la classification traditionnelle.

FINIR	VIVRE	PARTIR	JOINDRE	SAVOIR
vu nu il il ty ʒə	vu nu il il ty ʒə	vu nu il il ty ʒə	vu nu il il ty ʒə	vu nu il il ty ʒə
finis fini	viv vi	part par	ʒwap ʒwɛ	sav sɛ
e õ	e õ	e õ	e õ	e õ

BOIRE ²⁴	DEVOIR	VOULOIR	PRENDRE	VENIR
vu nu il il ty ʒə	vu nu il il ty ʒə	vu nu il il ty ʒə	vu nu il il ty ʒə	vu nu il il ty ʒə
byv bwav bwa	dəv dwav dwa	vul vœl vø	pɛən pɛən pɛā	vən vjən vjɛ
e õ	e õ	e õ	e õ	e õ

Tableau 18 – Quelques exemples de présentation de la conjugaison orale du Présent

La présentation suivante de l'ensemble du paradigme des tiroirs verbaux courants mettrait rapidement en évidence la distribution des bases orales:

	Pr		Imp		Pr S		Fut		Cond	
vu	byv	E	byv	je	byv	je	bwa	ʔe	bwa	ʔje
nu		õ		jõ		jõ		ʔõ		ʔjõ
il	bwav			E	bwav					ʔE
il	bwa							ʔa		
ty										
ʒə								ʔE		

Tableau 19 – La distribution des trois bases de *boire* aux cinq tiroirs verbaux simples usuels

Conclusion

En permettant une visualisation immédiate de la répartition des bases au travers de l'ensemble du paradigme, une telle présentation réduit considérablement la complexification résultant de la conjugaison orthographique. Toutefois, aussi anciennes soient les avancées linguistiques autorisant ce genre d'approche, force est de constater qu'elles demeurent en grande partie ignorées dans l'enseignement/apprentissage du FLE (cf. Germain & Séguin, 1998: 10).

²⁴ Sur les 55 verbes à trois bases – conjugués selon 7 verbes modèles: *boire, devoir, mouvoir, pouvoir, prendre, venir, vouloir* –, 15, à l'image de *boire*, mettent en partie en œuvre une opposition pluriel/singulier que cette disposition rend plus immédiate /bwav/-/bwa/, /dwav/-/dwa/, etc.

Pourquoi? La principale raison réside à notre avis dans la méconnaissance de la transcription phonétique, tant chez les apprenants que chez les enseignants²⁵. Aussi, serait-il naïf de proposer l'adoption d'une telle approche sans, au préalable, initier les apprenants (et les enseignants) à la lecture et l'écriture des symboles phonétiques. Beaucoup objecteront qu'une telle pratique constitue un surcroît de travail pour les débutants. Sur le court terme certes, mais l'apprentissage d'une langue étrangère s'envisage sur la durée. En dépit des premières réticences, la transcription phonétique se révèle un outil précieux d'apprentissage, qu'il est possible d'aborder de façon ludique et constructive (mots croisés phonétiques, le mot phonétique le plus long, etc.). Elle permet à la fois de donner une vision plus juste et moins fugitive de l'oral, tout en servant de passerelle vers l'orthographe²⁶.

Par ailleurs, la présentation inversée des personnes préconisée ici fait écho à d'autres phénomènes de la morphologie du français. La différence pluriel/singulier et la stratégie de soustraction à laquelle elle donne lieu (/sât/-/sâ/ *sentent*, *sent*) se retrouve également dans la différence féminin/masculin de certains adjectifs et de certains noms: /lât/-/lâ/ (*lente-lent*), /ɛzidât/-/ɛzidâ/ (*résidente*, *résident*), etc. (cf. Blanche-Benveniste, 1997: 143s; Blanche-Benveniste, 2010: 60s). En somme, contrairement à ce qu'a toujours imposé la tradition lexicographique ou grammaticale, le pluriel du verbe²⁷, et le féminin des noms et adjectifs s'avèrent préférables comme formes orales à mémoriser. Reste à savoir si dans l'apprentissage de la morphologie, l'efficacité sur le long terme l'emportera sur l'inertie de la tradition inspirée par l'enseignement du FLM, lequel n'a nul besoin de se préoccuper des difficultés de nos apprenants de FLE...

²⁵ En dépit des remarques de Catach (2004: 108): "L'API est absolument indispensable aujourd'hui, ne serait-ce que pour l'apprentissage des langues étrangères".

²⁶ En définitive, 17 symboles (/y/, /u/, /e/, /ɛ/, /ø/, /œ/, /ə/, /ɔ/, /õ/, /ã/, /ẽ/, /ʃ/, /ʒ/, /ɲ/, /ʎ/, /j/, /ɥ/) doivent être appris, soit parce qu'ils restituent des phonèmes différents de ce que la lecture des lettres imposerait normalement (par ex. /y/ ne se lit pas /i/, comme l'orthographe de *mythe* le laisse supposer), soit qu'ils sont absents de l'alphabet latin /ɛ/, /ø/, etc.

²⁷ "The signans of the plural tends to echo the meaning of a numeral increment by an increased length of the form" (Jakobson, 1965: 30). En d'autres termes, iconiquement, on désigne davantage d'entités avec des bases longues (pluriel), qu'avec des courtes (singulier).

Bibliographie

- Abry, D. & Chalaron, M.-L. (1997): La grammaire des premiers temps. Volume 1. Grenoble (Presses Universitaires de Grenoble).
- Arrivé, M.D. (2006): *Bescherelle. La conjugaison pour tous*. Paris (Hatier).
- Besse, H. & Porquier, R. (1984): *Grammaire et didactique des langues*. Paris (Hatier/Didier).
- Blanche-Benveniste, C. (1997): *Approches de la langue parlée en français*. Gap/Paris (Ophrys).
- Blanche-Benveniste, C. (2010): *Le français. Usages de la langue parlée*. Leuven/Paris (Peeters).
- Bonami, O. & Boyé, G. (2003): Supplétion et classes flexionnelles. In: *Langages*, 152, 102-126.
- Bonami, O. & Boyé, G. (2007): Remarques sur les bases de la conjugaison. In: Delais-Roussarie, E. & Labrune, L. (éds.): *Des sons et des sens*. Paris (Hermès Sciences), 77-90.
- Boyé, G. (2000): Problèmes de morpho-phonologie verbale en français, en espagnol et en italien, Université Paris VII. 503.
- Brissaud, C., Chevrot, J.-P. & Lefrançois, P. (2006): Les formes verbales homophones en /E/ entre 8 et 15 ans: contraintes et conflits dans la construction des savoirs sur une difficulté orthographique majeure du français. In: *Langue française*, 151, 74-93.
- Carelli, É., Fournier, G., Fuchs, M., Korach, D., Lancina, M. & Sabre, R. (1995): *Robert & Nathan. Conjugaison*. Paris (Nathan).
- Catach, N. (1995): *L'orthographe française*. Paris (Nathan).
- Catach, N. (2004): *L'orthographe*. Paris (Presses Universitaires de France (coll. Que sais-je?)).
- Csécsy, M. (1968): *De la linguistique à la pédagogie. Le verbe français*. Paris (Hachette/Larousse).
- Detey, S., Durand, J., Laks, B. & Lyche, C. (éds.) (2010): *Les variétés du français parlé dans l'espace francophone. Ressources pour l'enseignement*. Paris (Ophrys).
- Deutscher, G. (2009): "Overall complexity": a wild goose chase? In: Sampson, G., *et al.* (eds.): *Language complexity as an evolving variable*. Oxford (Oxford University Press), 243-251.
- Dubois, J. (1967): *Grammaire structurale du français. Le verbe*. Paris (Larousse).
- Éluerd, R. (2009): *Littre Conjugaison*. Paris (Garnier).
- Gaatone, D. (2001): De quelques principes de base pour la description de la morphologie verbale en français. In: Kronning, H., *et al.* (éds.): *Langage et référence. Mélanges offerts à Kerstin Jonasson à l'occasion de ses soixante ans*. Uppsala (Ubsaliensis Academiae), 213-222.
- Germain, C. & Séguin, H. (1998): *Le point sur la grammaire*. Paris (Cle International).
- Gougenheim, G., Michéa, R., Rivenc, P. & Sauvageot, A. (1964): *L'élaboration du français fondamental: étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Paris (Didier).
- Jaffré, J.-P. (2005): The Orthography of French. In: *Educational Studies in Language and Literature*, 5, 353-364.
- Jakobson, R. (1965): Quest for the Essence of Language. In: *Diogenes*, 13-51, 21-37.
- Katz, L. & Frost, R. (1992): The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis. In: Frost, R. & Katz, L. (eds.): *Orthography, phonology, morphology, and meaning*. Amsterdam/New York (North-Holland), 67-84.
- Kilani-Schoch, M. & Dressler, W.U. (2005): *Morphologie naturelle et flexion du verbe français*. Tübingen (Gunter Narr).
- Le Goffic, P. (1997): *Les formes conjuguées du verbe français. Oral et écrit*. Gap/Paris (Ophrys).
- Le Lay, Y. (1995): *Larousse Conjugaison*. Paris (Larousse).
- Léon, P. (2007): *Phonétisme et prononciations du français*. Paris (Armand Colin).
- Mahmoudian, M. (1976): *Pour enseigner le français. Présentation fonctionnelle de la langue*. Paris (Presses universitaires de France).
- Maingueneau, D. (1999): *L'énonciation en linguistique française*. Paris (Hachette).

- Marty, F. (1971): Les formes du verbe en français parlé. In: Rigault, A. (éd.): La grammaire du français parlé. Paris (Hachette), 105-117.
- Mayer, E. (1969): Stems in the French Verb. In: The French Review, 42-4, 555-568.
- Morin, Y.C. (2007): Les yods des terminaisons *-ions* et *-iez* de l'imparfait de l'indicatif et du présent du subjonctif des verbes français. In: Delais-Roussarie, E. & Labrune, L. (éds.): Des sons et des sens. Paris (Lavoisier), 124-140.
- Morin, Y.C. (2008): Les yods fluctuants dans la morphologie du verbe français. In: Fradin, B. (éd.): La Raison morphologique. Hommage à la mémoire de Danielle Corbin. Amsterdam/Philadelphia (John Benjamins), 133-154.
- New, B., Brysbaert, M., Veronis, J. & Pallier, C. (2007): The use of film subtitles to estimate word frequencies. In: Applied Psycholinguistics, 28, 661-677.
- Nichols, J. (2009): Linguistic complexity: a comprehensive definition and survey. In: Sampson, G., *et al.* (eds.): Language complexity as an evolving variable. Oxford (Oxford University Press), 110-125.
- Pinchon, J. & Couté, B. (1981): Le système verbal du français. Paris (Nathan).
- Pouradier Duteil, F. (1997): Le verbe français en conjugaison orale. Francfort/Main (Peter Lang).
- Riegel, M., Pellat, J.-C. & Rioul, R. (2009): Grammaire méthodique du français. Paris (Presses Universitaires de France).
- Séguin, H. (1989): Pour une taxinomie de la morphologie verbale en didactique du français. In: Études de linguistique appliquée, 75, 124-137.
- Thimonnier, R. (1967): Le Système graphique du français. Paris (Plon).
- Touratier, C. (1996): Le système verbal français. Paris (Armand Colin).
- Wagner, R.L. & Pinchon, J. (1991): Grammaire du français classique et moderne. Paris (Hachette).

Pratiques orthographiques en français d'étudiants étrangers: le cas d'étudiants hispano-américains et afghans

Jeanne GONAC'H & Clara MORTAMET

IRED, Université de Rouen & UFR Lettres et sciences humaines, Université de Rouen

In this article, we analyze and compare French spelling errors produced by two different groups of university students who have learned French in a country where French is a foreign language (FFL): students from Central and South American countries (Salvador, Peru and Argentina) and Afghanistan. Our corpus is based on errors found in papers of graduate distance-learning students, most of whom are planning on becoming French teachers in their countries. Our aim is to identify what errors are specific to FFL university students and eventually what errors are specific to the two different groups. The first results of our study confirm tendencies observed on previous research and prove that there exist specific types of errors for FFL learners. In the end, we will also discuss the limit of error categories.

Introduction

S'interroger sur l'élaboration de méthodes d'enseignement spécifiques pour pallier les problèmes d'orthographe des apprenants en contexte FLE implique en amont d'identifier les problèmes particuliers que l'orthographe française pose aux apprenants allophones. Or, les études qui ont inventorié les erreurs spécifiques à ces apprenants d'une part sont peu nombreuses et d'autre part montrent que la majorité des erreurs des apprenants FLE, si elles sont plus nombreuses que celles de leurs pairs natifs, ne sont pas pour autant toujours différentes de celles de ces derniers (Amokrane & Brissaud, *à paraître*). Cette ressemblance entre les erreurs des natifs et des non-natifs dans des copies universitaires a été confortée dans une étude contrastive que nous avons menée récemment à partir de copies d'étudiants éthiopiens et mahorais (Mortamet & Gonac'h, 2011), étude qui s'est ajoutée à l'expérience que nous avons de l'analyse de l'orthographe des natifs monolingues et bilingues (Mortamet, 2003; Gonac'h, 2009).

Notre contribution vise ainsi à compléter l'analyse des zones orthographiques problématiques pour des apprenants en contexte FLE. Pour ce faire, nous avons choisi d'analyser les erreurs produites dans des copies d'examen rédigées par des étudiants de 3^{ème} année de licence inscrits en sciences du langage en télé-enseignement à l'université de Rouen. En effet, par le biais de ce centre, nous avons accès à des dizaines

de copies du monde entier. Ces adultes en formation sont répartis sur tous les continents, très majoritairement dans les pays du Sud. Beaucoup de ces étudiants n'ont pas le français pour langue première, mais l'enseignent ou souhaitent l'enseigner dans leur pays. Ce corpus offre la possibilité de comparer des productions émanant de contextes d'apprentissage très divers. Dans le cadre de cette étude, nous explorons deux contextes de français langue étrangère: l'Amérique hispanophone d'une part, et l'Afghanistan d'autre part.

Dans un premier temps, nous ferons le point des travaux, relevant tant de la sociolinguistique que de la didactique, qui nous ont inspirées pour penser la question des erreurs d'orthographe chez les non-natifs. Nous présenterons ensuite la méthodologie de recueil et d'analyse des productions. Nous analyserons enfin les principales erreurs observées dans notre corpus, qu'elles relèvent de la fonction morphographique, visuographique, phonographique, ou qu'elles fassent l'objet d'interprétations mixtes.

1. Les erreurs d'orthographe des non-natifs comme objet de recherche

1.1 Un objet au croisement de plusieurs domaines de recherche

La question des erreurs d'orthographe des non-natifs nécessite une expertise à la fois sur l'orthographe, sur le plurilinguisme et sur l'acquisition-apprentissage des langues.

En ce qui concerne l'analyse des erreurs d'orthographe, nous nous appuyons d'abord sur les travaux de Catach (1986, 1998), mais aussi sur ceux de Brissaud (2006), de Honvault (dir., 2006), de Lucci & Millet (1994), de Manesse & Cogis (2007). Ces études, s'appuyant toutes sur l'analyse de l'orthographe française de Catach (phonographie, morphographie, idéographie, logographie), qu'elles affinent au passage, présentent l'intérêt de décrire des pratiques orthographiques dans des activités diverses (des dictées comme dans Manesse & Cogis, des écrits ordinaires comme dans Lucci & Millet, des écrits scolaires dans Brissaud), par des locuteurs divers (des élèves dans Manesse & Cogis, Brissaud, des secrétaires et des étudiants dans Lucci & Millet). Elles donnent également à voir des pratiques issues de situations plus ou moins formelles, plus ou moins publiques (des brouillons dans Lucci & Millet, des écrits scolaires dans Brissaud, Manesse & Cogis). Mais elles nous ont surtout servi ici à établir notre grille d'analyse des erreurs, et à comparer les erreurs de notre corpus avec celles rencontrées dans ces corpus de natifs. Elles ont enfin alimenté notre réflexion plus générale sur l'orthographe du français et les pratiques orthographiques des francophones (Honvault (dir.), 2006).

En ce qui concerne les questions de plurilinguisme et de didactique des langues, nous nous appuyons essentiellement sur les travaux développés par Porquier & Py (2005), Selinker (1994), Tarone (2005), et ceux menés dans le cadre du CECR. Les réflexions nées autour de la notion de *compétence plurilingue* (Coste, Moore & Zarate, 1997: 105-106) ont ouvert la voie à une nouvelle conceptualisation des liens entre compétence et pratique des langues présentes dans le répertoire sans faire intervenir l'idée de l'influence négative d'une langue sur l'autre. Coste (2002: 120) souligne ainsi que les compétences de toutes les ressources langagières présentes dans le répertoire sont mises en relation. Il fait l'hypothèse que les *éventuelles* différences entre les compétences des non-natifs et des natifs sont liées à la "circulation interlinguistique" des répertoires. Avec Coste, nous faisons donc l'hypothèse que les langues présentes dans le répertoire des non-natifs peuvent affecter leur orthographe, et donner lieu à des erreurs "différentes" de celles des natifs – la "différence" pouvant porter sur les types d'erreurs, ou sur la proportion de chaque type d'erreurs dans l'ensemble des erreurs. Cette hypothèse implique aussi que les pratiques orthographiques des apprenants peuvent être affectées différemment lorsque les répertoires langagiers sont distincts. C'est ce que nous nous attacherons à regarder ici en comparant les pratiques d'étudiants afghans et hispano-américains en français.

1.2 *Quelques particularités des non-natifs*

Dans une précédente recherche (Mortamet & Gonac'h, 2011), nous avons comparé des copies rédigées dans un contexte FLE (l'Ethiopie) et dans un contexte FLS (Mayotte) du point de vue de l'orthographe. Notre hypothèse de départ était alors que les étudiants en contexte FLE produiraient des erreurs différentes de celles des étudiants en contexte FLS. Nos résultats ont plutôt infirmé cette hypothèse: il apparaît des erreurs spécifiques dans ces deux contextes, qui les différencient des natifs mais pas entre eux. En particulier, nous avons montré que les non-natifs, quel que soit leur contexte d'apprentissage (FLE et FLS), produisaient *en proportion* davantage d'erreurs d'accord en genre que ce que l'on rencontre dans des corpus de natifs.

Dans une étude portant sur des natifs, Brissaud *et alii* (2006) indiquent par exemple que les erreurs d'accord en genre figurent parmi les plus fréquentes, mais qu'elles ne sont pas plus fréquentes que celles portant sur les accords en nombre ou sur les cas d'homophonie verbale. Dans l'étude de Manesse & Cogis (2007: 111ss.), qui porte également sur un corpus de natifs, les accords en genre apparaissent même quelque peu plus réussis que les accords en nombre.

Dans notre corpus d'erreurs produits par les étudiants éthiopiens et mahorais, les erreurs liées au genre sont non seulement plus nombreuses mais aussi différentes de celles des natifs. De plus, elles ne relèvent pas uniquement de l'orthographe (*un problème majeure, la méthodologie audio-visuel*) mais aussi, bien souvent de la morphologie (*la problème, un public instruite*, etc.). A partir de ces résultats, nous avons fait l'hypothèse que l'apprentissage du genre grammatical en français, que ce soit par l'orthographe ou par la morphologie, posait davantage problème aux étudiants non-natifs qu'à leurs pairs natifs notamment parce que le marquage du genre grammatical dans leurs langues premières n'existe pas ou ne fonctionne pas de la même façon.

Nous appuyant sur cette hypothèse de circulation interlinguistique, nous avons voulu ici prolonger ce travail dans cette direction en comparant les erreurs produites par des étudiants dont la langue première ou principale est typologiquement proche du français (l'espagnol) avec celles des étudiants dont la langue première ou principale est très éloignée du français (le dari).

2. Méthodologie

2.1 *Choix du corpus*

Nous avons constitué deux groupes d'étudiants: huit étudiants afghans d'une part et huit étudiants hispano-américains (6 Salvadoriens, 1 Péruvienne et 1 Argentine) d'autre part. Parmi les étudiants afghans, cinq sont des femmes; ils ont autour de 30 ans à l'exception d'une personne qui en a 60. Parmi les étudiants hispano-américains, sept sont des femmes. Ils sont âgés entre 25 et 30 ans et une femme a plus de 60 ans.

Pour chaque étudiant, nous avons analysé cinq de leurs copies d'examens de troisième année de licence de Sciences du langage parcours FLE. Nous avons toujours choisi des copies rédigées de façon libre (dissertations, commentaires de texte, etc.).

2.2 *Méthodologie d'analyse*

Pour chaque étudiant, nous avons relevé l'ensemble des erreurs produites et les avons ensuite classées selon différentes catégories¹. Dans un premier temps, nous analyserons ici plus particulièrement les catégories suivantes: les problèmes liés à l'accord en genre qui relèvent de l'orthographe (*un nouvelle emploi*) et de la morphologie (*son rigueur*)² (§3);

¹ Pour une présentation complète des catégories, voir Mortamet & Gonac'h, 2011.

² Ces deux types d'erreurs étant très représentés (tous les étudiants en font) et répétés dans les mêmes copies, il nous a semblé que ces formes devaient être analysées ensemble. Pour

les problèmes d'accord en nombre qui relèvent de l'orthographe (*ces communication*) (§4); les erreurs liées aux lettres hors système (*auteur, manière*) (§5); les problèmes qui relèvent de la phonographie (*contrasens, distiné*) (§6). Nous analyserons aussi les erreurs qui peuvent avoir des origines différentes et qui peuvent donc être interprétées de deux façons différentes (§7).

Nous n'analyserons pas les écarts portant sur les diacritiques et les auxiliaires d'écriture (accents, traits d'union, apostrophes), les césures, les "ratés graphiques", ni sur la ponctuation, ces zones d'écarts soulevant chacune des questions particulières, difficiles à articuler avec celles présentées ici³. Par ailleurs, nous écarterons aussi certaines catégories très faiblement représentées dans nos copies. Il s'agit notamment des accords sur les verbes conjugués, des anglicismes, des erreurs de morphographie lexicale (*aspet*), des erreurs sur les morphogrammes (*judicière*).

Une fois toutes ces formes relevées et classées, nous avons compté les occurrences dans chaque catégorie pour chaque étudiant (toutes copies confondues). Cela nous a permis d'établir des diagnostics de chaque étudiant, de regarder les écarts les plus fréquemment rencontrés chez chacun d'entre eux, ceux qu'ils avaient en commun, ceux qui les distinguaient. Par contre, nous n'avons pas fait la somme par groupe des erreurs de chaque type. Nous avons en effet voulu adopter une analyse qualitative, en établissant un diagnostic par étudiant, et en comparant ensuite les deux ensembles de diagnostics. La comparaison de nos deux populations, que nous présentons maintenant, ne s'appuie donc pas sur un diagnostic global de chaque groupe d'étudiants, mais sur un ensemble d'évaluations individuelles qui composent chaque groupe.

la même raison, nous avons analysé ensemble tous les problèmes de genre, qu'ils portent sur les participes passés, les adjectifs, les attributs ou les noms. Nous avons fait la même chose avec les erreurs concernant le nombre.

³ Les diacritiques et auxiliaires d'écriture ne font pas partie, ou à des degrés divers, du plurisystème graphique: si les accents aigus et graves sur le <e> et le tréma présentent pour certains encore une dimension phonographique, elle est dans certains cas discutée, et ne vaut plus pour les autres diacritiques et auxiliaires que sont les traits d'union, les apostrophes, l'accent circonflexe. La question de la césure, absente de notre corpus du fait du niveau avancé de nos scripteurs, pose des questions diverses, entre morphographie et visuographie. Les "ratés graphiques" (répétitions de syllabes comme dans *institutution*, omissions de syllabes dans *instution*), relativement nombreux, sont pour l'essentiel dus à la contrainte physique de l'écriture manuscrite et difficilement interprétables, *a priori* du moins, comme des erreurs d'orthographe. La ponctuation enfin relève d'un autre niveau, celui de la phrase et du texte.

3. Les erreurs portant sur le genre

Dans l'ensemble du corpus, les erreurs liées à l'accord en genre qui relèvent de l'orthographe et de la morphologie sont un peu plus nombreuses que les erreurs liées à l'accord en nombre. Il faut tout de suite souligner que les étudiants afghans contribuent largement à ce volume et que parmi eux, certains produisent largement plus d'erreurs que d'autres (l'un d'entre eux produit plus de 30 erreurs alors qu'on en relève 140 dans l'ensemble des copies de ce groupe). Autrement dit, sur ce point, l'hétérogénéité interne au groupe est plus forte chez les étudiants afghans que chez les étudiants hispano-américains.

Il n'en reste pas moins que ces résultats rejoignent ce que nous avons déjà constaté, à savoir que les variations liées au genre sont les plus nombreuses. Elles sont bel et bien surreprésentées dans les productions des étudiants dont le français n'est pas la langue première par rapport aux productions des natifs (cf. *supra*, §1.2).

3.1 Particularités des étudiants afghans

Les étudiants afghans se distinguent particulièrement par leurs erreurs portant sur le genre morphologique, catégorie à laquelle ils contribuent beaucoup plus que les étudiants hispano-américains. Nous avons repéré dans cette catégorie plusieurs cas de figure: soit les étudiants utilisent les marques du masculin alors que les marques du féminin sont attendues soit l'inverse; les marques du masculin peuvent être portées par le déterminant: *son rigueur*; par l'adjectif: *la compétence communicatif* et peuvent aussi être multiples comme dans: *le revu* pour *la revue*. Les marques du masculin portées par le déterminant sont les erreurs les plus fréquentes et les plus partagées chez les étudiants afghans (7 étudiants sur 8). Le nombre d'erreurs de ce type varie de façon importante à l'intérieur du groupe puisque l'un d'entre eux en fait dix, et que cinq autres n'en font qu'une. Les marques du masculin (ou l'absence de marque du féminin) sur l'adjectif et les marques "multiples" concernent quatre étudiants.

Parmi ces erreurs, certaines indiquent que des terminaisons nominales régulières, qui sont plutôt typiques d'un genre, ne sont pas reconnues par les étudiants. C'est le cas par exemple des terminaisons du féminin *-ion* (mission), *-ologie* (méthodologie), *-ure* (candidature).

Les marques du féminin sont aussi souvent utilisées quand celles du masculin sont attendues. Alors que les marques du masculin portaient plus souvent sur le déterminant seulement (*le façon*), les marques du féminin les plus partagées sont celles qui portent uniquement sur l'adjectif (*diverses domaines*). Ensuite, on compte les erreurs qui impliquent l'ajout

des marques du féminin à la fois sur les déterminants et les adjectifs (*une phénomène sociale*) que cinq scripteurs différents produisent. Viennent ensuite les erreurs sur le déterminant seul (*la couloir*), et enfin les marques sur les participes passés, qu'un seul étudiant produit (*l'apprentissage s'est développée*).

Nous avons distingué les marques du féminin et du masculin afin de faire émerger des tendances: soit à "masculiniser" ou à l'inverse à "féminiser". Toutefois, cette distinction ne s'est avérée pertinente que pour un étudiant afghan qui a tendance à féminiser. Pour tous les autres, on ne repère aucune tendance particulière.

Comme signalé plus haut, les confusions sur le genre au niveau morphologique sont beaucoup moins nombreuses pour les étudiants hispano-américains. Seulement la moitié des étudiants produit des erreurs de ce type et parmi eux, une étudiante – péruvienne – se détache en en produisant beaucoup plus que les autres. Les erreurs ne sont pas liées à un calque de l'espagnol à l'exception de "*La petite chaperon rouge*" (*Caperucita Roja, féminin en espagnol*). Ainsi, par exemple, un étudiant écrit "*la texte*" alors qu'en espagnol "texte" est aussi masculin (*el texto*).

La comparaison des deux contextes d'apprentissage révèle que la question du genre ne pose pas les mêmes problèmes aux étudiants afghans et aux étudiants hispano-américains. Alors que les étudiants afghans utilisent tous, souvent et à différents niveaux, des marques du féminin ou du masculin quand l'inverse est attendu, seules quelques erreurs apparaissent dans le corpus de copies des Hispano-américains. Le marquage du genre grammatical dans les langues premières peut être un des motifs d'explication de cette différence entre les deux groupes. Ainsi, en dari, langue que tous nos étudiants pratiquent, le genre grammatical n'est pas marqué, alors qu'en espagnol il l'est. Les quelques erreurs produites par les Hispano-américains peuvent être liées au fait que les genres grammaticaux du français et de l'espagnol ne correspondent pas toujours. Pour les Afghans, on peut penser que les erreurs sont liées au surcoût cognitif que représente l'apprentissage du genre grammatical. Ainsi, on peut même penser que c'est la logique de la "loterie" qui domine dans leurs copies lorsqu'ils ne connaissent pas le genre. Si cette interprétation est juste, il y a fort à parier que chez les Afghans étudiés, les méconnaissances du genre des noms utilisés sont plus nombreuses que celles qui apparaissent ici: si les cas de loterie malchanceuse sont remarquables, ceux de loterie chanceuse se confondent avec les formes issues d'une bonne maîtrise du genre.

3.2 *Erreurs communes aux deux groupes*

Pour les étudiants afghans comme pour les étudiants hispano-américains, les problèmes d'accord orthographique liés au genre portent essentiellement sur l'absence d'accord au féminin de l'adjectif (*une notion flou; des études supérieurs*) et aussi et non moins souvent sur l'ajout de lettres qui correspondent grammaticalement au féminin sur l'adjectif (*un nouvelle emploi; les mots inconnues*). Là encore, on peut trouver dans les langues premières une explication à cette convergence de pratiques. Si les hispanophones rejoignent ici les Afghans, c'est que malgré leur usage d'un genre grammatical, celui-ci ne se manifeste jamais uniquement orthographiquement: quand le genre n'est pas porté par l'adjectif à l'oral (comme c'est le cas pour *negro/negra*), la forme écrite est épiciène (*suave/suave*) mais on ne connaît pas de cas comme *noir/noire*. Les erreurs sur les participes passés et les attributs qui ont été intégrées à cette catégorie ne sont pas nombreuses (*elle est basé*) et restent donc marginales. Un étudiant hispano-américain se distingue néanmoins ici car, contrairement aux autres, ses erreurs portent essentiellement sur les participes passés. Pour autant, cette particularité ne signale pas pour nous une incompétence particulière de cet étudiant sur la question des participes passés. En effet, autant il est difficile de ne pas employer de noms, et donc de masquer la non-reconnaissance de leur genre, autant l'emploi des participes passés peut être évité ou du moins limité dans des textes tels que ceux que nous étudions. Cet étudiant fait davantage d'erreurs car il utilise davantage de participes passés, et sa particularité tient surtout à sa prise de risques. Enfin, sauf exception, on ne relève pas non plus de tendance à masculiniser ou à féminiser tant au niveau des groupes qu'au niveau des étudiants.

3.3 *Bilan*

Pour les deux groupes, les erreurs de genre au niveau morphologique et au niveau orthographique portent en majorité sur le syntagme nominal. Le genre est une des difficultés majeures pour les étudiants afghans, au niveau morphologique et au niveau orthographique. Les étudiants hispano-américains rejoignent les étudiants afghans uniquement sur les erreurs au niveau orthographique. Par rapport aux hypothèses qui pourraient éclairer ce résultat, la question du genre grammatical dans les langues premières nous semble être une des plus pertinentes.

4. Le nombre

La question des accords en nombre relève essentiellement de l'orthographe (*les langue*) et seulement quelques erreurs marginales sont

d'ordre morphologique (*le document authentique oraux*); nous laisserons donc ces dernières de côté. Les erreurs liées à l'accord en nombre sont la 2^{ème} catégorie la plus nombreuse pour l'ensemble des étudiants des deux groupes.

4.1 *Des accords en nombre désordonnés chez les étudiants afghans*

Parmi les différents types d'erreurs portant sur le nombre, nous n'avons pas été étonnées de l'absence de <s> sur le nom quand celui-ci est précédé d'un déterminant au pluriel (*les texte*) ni de l'absence de <s> sur l'adjectif (*des documents authentique*) pas plus que de l'absence de <s> sur les deux (*des raison typique*). En revanche, l'absence de <s> sur le nom alors que le déterminant et l'adjectif sont marqués au pluriel (*les texte littéraires*) nous a plus surprises. Ce type d'erreurs est en effet le deuxième le plus partagé par les étudiants afghans (cinq d'entre eux en produisent) et il n'est donc pas marginal. Ce non-marquage du nom pris "en sandwich" entre deux formes au pluriel suggère deux interprétations possibles. Ces erreurs peuvent être mises en relation avec le fait que les compétences des étudiants ne sont pas systématisées sur ce point: l'apprentissage est en cours, et l'attention est portée sur la marque du pluriel du déterminant et de l'adjectif en priorité. Ces erreurs peuvent aussi être éclairées par le marquage du pluriel en dari. En effet, en dari, à l'oral et à l'écrit, le pluriel est porté par le déterminant uniquement (des langues étrangères: *zaban hay kharji*; une langue étrangère: *yak zaban kharji*). Au contraire, en français, le pluriel est marqué par l'ensemble des composants du syntagme nominal à l'écrit, mais seulement par le déterminant à l'oral, à l'exception des pluriels dits "irréguliers" (*des chevaux*) et de la liaison entre l'adjectif antéposé et le nom (*des petits amis*). Ces différences de marquage peuvent être sources de confusion pour l'apprenant qui n'est pas habitué à marquer le nombre par un morphogramme, mais toujours par un phonogramme.

Du point de vue des étudiants (et non plus des types d'erreurs), on remarque que les quatre étudiants qui produisent le plus d'erreurs d'accord en nombre font aussi partie de ceux qui font le plus d'erreurs sur le genre. Généralement, la tendance au niveau de la fréquence des erreurs de genre rejoint celle des erreurs de nombre.

4.2 *Les erreurs attendues des étudiants hispano-américains*

Les Hispano-américains produisent beaucoup moins d'erreurs et moins d'erreurs de types différents que les étudiants afghans: le type d'erreur le plus fréquent est, de même que pour les Afghans, celui de "*ces communication*". On ne rencontre pas non plus beaucoup d'erreurs sur les

participes passés et les attributs; elles se concentrent essentiellement sur le syntagme nominal.

Dans l'ensemble ces erreurs sont marginales dans ce corpus. Un étudiant se distingue en produisant davantage d'erreurs de ce type. Il faisait aussi partie de ceux qui produisaient le plus d'erreurs sur le genre au niveau orthographique. De la même façon que dans le corpus afghan, la tendance au niveau de la fréquence des erreurs sur le nombre rejoint donc celle des erreurs sur le genre.

5. Les erreurs visuographiques

Parmi les erreurs visuographiques, nous analyserons ici uniquement celles qui concernent les lettres hors système. Il s'agit toujours d'ajouts et d'omissions de lettres étymologiques et de consonnes doubles chez les Afghans et les Hispano-américains.

C'est la première catégorie d'erreurs qui n'oppose pas les deux groupes d'étudiants. On relève une trentaine d'erreurs pour chaque groupe: ajout ou omission de <h>: *autheur*, *authentique*; ajout ou omission de <n>: *mentioner*, *mannière*; etc. Nous avons même relevé une erreur identique dans les deux groupes: *synonime*. De manière générale, il n'existe pas pour chaque étudiant de tendance à ajouter ou à omettre des lettres étymologiques ou des consonnes doubles.

Sur ce point, nous n'avons pas relevé dans notre corpus plus d'erreurs ou des erreurs différentes de celles produites par des natifs (telles qu'elles apparaissent dans les études déjà citées de Brissaud *et alii*, 2006, Manesse & Cogis, 2007, Jaffré & Bollengier, 1995). Autrement dit, les lettres hors système, qui constituent une des spécificités de l'orthographe française, ont été dans l'ensemble rapidement intégrées au système orthographique des apprenants FLE. Notons au passage que ce résultat nuance sinon contredit l'opinion de certains spécialistes de l'orthographe qui appellent à une réforme de l'orthographe sur ces lettres hors système en priorité. Finalement, malgré leur caractère non motivé, elles posent plutôt moins de difficultés, y compris chez les apprenants de FLE, que la question des accords ou de la transcription des phonèmes.

6. Les erreurs phonographiques

Dans cette catégorie, nous avons classé les erreurs du type: *répétation* pour *répétition*.

6.1 *Des phonogrammes vocaliques problématiques pour les étudiants afghans*

Les étudiants afghans produisent en majorité des erreurs au niveau des graphèmes chargés de transcrire les phonèmes vocaliques. Par exemple, on compte pour le seul phonème /E/⁴ quatre correspondants graphiques différents ("*des interjonction*", "*portinent*", "*logiciale*", "*distiné*"), pour le phonème /O/, deux graphèmes différents ("*communité*"; "*emboucher*") ainsi que pour le phonème /i/ ("*déchotomie*", "*répétation*"), etc. Sept des huit étudiants afghans produisent au moins une erreur de type phonographique. La répartition de ces erreurs est très inégale au sein du groupe. Ainsi, l'un d'entre eux produit un quart du total des erreurs de ce type.

Ce type d'erreur est le plus attendu dans les copies de locuteurs non-natifs, et on les trouve d'ailleurs aussi dans les copies des Ethiopiens (Mortamet & Gonac'h, 2011), mais surtout dans celles des Maghrébins analysées par Amokrane et Brissaud (2011). Il s'agit d'une conséquence naturelle de la complexité du système vocalique du français.

6.2 *Quelques phonogrammes consonantiques chez les étudiants hispano-américains*

La zone phonographique ne pose pas de difficultés comparables aux étudiants hispano-américains. Et, contrairement au groupe des Afghans, ce sont davantage quelques graphèmes consonantiques qui leur posent problème, en particulier le phonème /s/ (*s'agissait*, *accessible*) et les <ss> pour *suffisant* par exemple. On peut ajouter que la voyelle nasale /ɛ̃/ est souvent traduite à l'écrit par le graphème <an> (*salvadorian*). Enfin, il apparaît nettement que la plupart des erreurs de type phonogrammique sont des calques de l'espagnol: *acent* (< *acento*); *contra* (< *contra*); *divide* (< *divide*); peuple (< *pueblo*). Ainsi, la parenté linguistique de l'espagnol avec le français, si elle a pu favoriser certaines compétences orthographiques, peut aussi, comme ici, freiner l'apprentissage de formes pourtant assez courantes.

7. Les erreurs donnant lieu à plusieurs interprétations

Il convient pour terminer de revenir sur les erreurs pouvant relever de plusieurs catégories, sans que l'on puisse à aucun moment déterminer si l'erreur relève de l'une ou de l'autre, c'est-à-dire d'un problème visuographique, phonographique ou encore morphographique. Ces erreurs

⁴ Nous prenons ici pour référence le système phonologique restreint, qui neutralise l'opposition /e/ /ɛ/ (/E/), et /o/ /ɔ/ (/O/).

ne constituent pas un bon indicateur de différences ou spécificités entre les groupes: elles ne permettent en effet ni de caractériser nos corpus – on en trouve probablement dans d'autres corpus, y compris chez des natifs – et elles ne permettent pas non plus de distinguer nos deux groupes, ni même de distinguer des étudiants à l'intérieur de chaque groupe. Pour autant, il nous semble important de les présenter et de les analyser ici, parce qu'elles interrogent les catégories établies, et qu'elles posent la question de l'interprétation des erreurs, préalable nécessaire mais pas toujours possible à leur remédiation.

7.1 *Visuographie et morphographie grammaticale*

Ces erreurs impliquent essentiellement deux lettres de notre système: le <e> et le <s> en position finale, qui assurent de nombreuses fonctions, comme pour ce qui nous intéresse ici une fonction morphologique – la marque du féminin (une amie), celle du pluriel (des copains) – et une fonction visuographique (le mois; la vie). Nous avons ainsi relevé une vingtaine d'erreurs sur le <e> et une dizaine d'erreurs sur le <s> assurant ces fonctions.

Prenons l'exemple de l'omission du <e> dans *mond*, *norm* et *etap*. Dans ce cas on peut dire que cette omission est liée à une méconnaissance des règles de combinaison et de position des lettres pour constituer les graphèmes. La règle en l'occurrence est que si le phonème consonantique est en position finale de mot, le graphème consonantique est accompagné d'un <e> diacritique qui le marque en tant que phonogramme; en l'absence de ce <e>, il est fréquent d'avoir affaire à un morphogramme. C'est le cas par exemple de *porte* / *port*. Cette règle vaut surtout pour les phonogrammes *t*, *d*, *s*, avec quelques exceptions, mais est beaucoup moins systématique pour *r* ou *l* (*jovial* mais *fidèle*). Ce type d'erreurs est souvent relevé chez les jeunes apprenants natifs qui ne maîtrisent pas encore ces règles. Elles sont plus rares chez les apprenants qui connaissent les règles et sont habitués à l'image de ces terminaisons de mots.

Pour les exemples d'omission tels que *écrire* et *facil*, qui, nous l'avons vu, concernent des phonogrammes qui fonctionnent moins systématiquement avec un <e> diacritique final, on peut penser que les étudiants font une analogie dans le premier cas avec des infinitifs comme *finir*, *partir* et dans le second cas avec des adjectifs terminés par <l> comme *amical*.

A l'inverse, lorsque l'étudiant ajoute des <e> quand ils ne sont pas nécessaires, on peut penser qu'il "surapplique" la règle de position précédente. C'est le cas en particulier des <e> diacritiques ajoutés après des consonnes liquides comme dans *erreure*, par analogie à *heure*. On peut aussi penser que les étudiants les ajoutent par analogie à ce qui se fait

pour certains noms d'animés, comme *professeure*. Dans ce cas ils fonctionnent comme un genre référentiel et non grammatical.

Nous retrouvons également des analogies de ce type quand les scripteurs omettent un <s> comme dans: *le moi* (quand *mois* est attendu) ou quand ils en ajoutent comme dans: *le soucis*. Il semble évident ici que l'existence de formes comme *souris* les rend familières à l'œil.

7.2 *Phonographie et morphographie grammaticale*

Une erreur telle que *la compréhension écrit* peut relever à la fois du cas précédent – méconnaissance des règles de combinaison graphique – et d'un cas d'omission de la marque du féminin sur un adjectif qualificatif non épïcène. Dans le second cas, il s'agit d'un problème de morphologie orthographique (§3).

7.3 *Phonographie et morphographie lexicale*

Nous avons relevé deux erreurs dans le corpus hispano-américain, *linguistique* pour linguistique et *briève* pour brève, qui nous font penser d'abord que l'image acoustique des mots n'est pas stable dans le répertoire de ces apprenants et que leur graphie est donc non maîtrisée. En même temps, on peut considérer que le raisonnement qu'ils ont adopté pour orthographier ces mots relève de la morphographie lexicale: *brièvement* – *briève* (en espagnol: *brevemente* – *breve*).

On le voit, ces derniers cas d'erreurs sont impossibles à interpréter de façon univoque. On ne peut se livrer à une interprétation, et donc tenter une remédiation ciblée, en l'absence d'un échange avec les scripteurs, ou du moins de tests spécifiques complémentaires. Ces erreurs étant par ailleurs assez rares, mais aussi combinées avec des erreurs relevant de l'une des interprétations possibles, on ne peut guère s'appuyer non plus sur des régularités de scripteurs.

8. Conclusion

Ce travail a permis de montrer qu'il existe bien des zones de l'orthographe qui sont moins bien maîtrisées par les étudiants FLE que par les étudiants natifs et qu'il convient, donc, de construire des méthodologies d'apprentissage spécifiques pour ces publics. L'analyse contrastive des copies des étudiants afghans et hispano-américains a d'abord permis de dégager des points communs à chacun des groupes: accords en genre et en nombre, ajout / omission de lettres hors système, choix de phonogrammes ne correspondant pas au phonème. Les problèmes d'accord en nombre ainsi que les erreurs liées aux lettres hors système ne sont pas spécifiques des non-natifs, on en relève également dans les copies des étudiants

natifs. Il nous semble que ces erreurs sont davantage à mettre en liaison avec la spécificité et la complexité de l'orthographe française qu'avec la question de la maîtrise du français par les scripteurs. Les erreurs liées à la transcription des phonèmes sont quant à elles spécifiques aux étudiants non-natifs. Contrairement aux autres, ces erreurs sont à mettre en lien avec la circulation interlinguistique des répertoires des apprenants.

L'analyse contrastive a aussi permis de dégager des spécificités de chacun des groupes. C'est à partir de ces spécificités que nous pouvons maintenant apporter des éléments de discussion à notre hypothèse initiale – celle selon laquelle la distance typologique des langues présentes dans le répertoire des apprenants a un effet sur les formes qu'ils produisent en français. A propos du genre morphologique, il apparaît bien qu'il fait l'objet de plus d'erreurs chez les étudiants afghans que chez les étudiants hispano-américains. On peut penser que l'apprentissage du genre grammatical pose davantage de problèmes aux étudiants afghans parce que le dari n'a pas de genre grammatical, contrairement au français et à l'espagnol. Et si les étudiants hispano-américains font parfois des erreurs sur certains mots qui ont le même genre grammatical en espagnol et en français, c'est parce que le genre grammatical des deux langues n'est pas systématiquement symétrique. Néanmoins, il nous semble bien que la proximité du français et de l'espagnol joue un rôle dans la "meilleure" maîtrise de l'orthographe des étudiants hispanophones. Il conviendra de confirmer ce résultat dans l'analyse de nouveaux corpus. Enfin, cette analyse a montré que certaines interprétations des écarts d'orthographe étaient nécessairement multiples, et qu'une remédiation en orthographe devait aussi tenir compte de la diversité des voies qui mènent aux mêmes erreurs.

Bibliographie

- Amokrane, S. & Brissaud, C. (à paraître): L'apprentissage de l'orthographe en FLM et FLE/FLS: entre convergences et spécificités. In: *Langues modernes*.
- Brissaud, C. (2006): Le cas des homophones. In: *Cahiers pédagogiques*, 440, 47-48.
- Brissaud, C., Chevrot, J-P. & Lefrançois, P. (2006): Les formes verbales homophones en /E/ entre 8 et 15 ans: contraintes et conflits dans la construction des savoirs sur une difficulté orthographique majeure du français. In: *Langue Française*, 151. Paris (Larousse), 74-93.
- Catach, N. (1986): *L'orthographe française. Traité théorique et pratique*. Paris (Nathan).
- Catach, N. (1998): *L'orthographe, Que sais-je?* Paris (PUF) 8e édition.
- Coste, D. (2002): Compétence à communiquer et compétence plurilingue. In: Castellotti, V. & Py, B. (éds): *La notion de compétence en langue, Notions en questions*, 6. Lyon (ENS éditions), 115-123.

- Coste, D., Moore, D. & Zarate, G. (1997): Compétence plurilingue et pluriculturelle. Vers un Cadre Européen Commun de référence pour l'enseignement et l'apprentissage des langues vivantes. Strasbourg (édition du Conseil de l'Europe).
- Gonac'h, J. (2009): Des variations orthographiques dans les productions de jeunes bilingues turc-français et turc-anglais. In: *Langage et société*, 130, 105-122.
- Honvault-Ducrocq, R. (dir.) (2006): *L'orthographe en questions*. Rouen (PUR).
- Jaffré, J-P. & Bollengier, F. (1995): Le lycée professionnel et l'orthographe. In: Anis, J. & Cusin-Berche, F. (éds.): *Difficultés linguistiques des jeunes en formation professionnelle courte*, numéro spécial Linx, 279-289.
- Lefrançois, P. (2001): Le point sur les transferts dans l'écriture en langue seconde. In: *Canadian Modern Language Review*, 58. Toronto (University of Toronto Press).
- Lucci, V. & Millet, A. (dirs.) (1994): *L'Orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*. Paris (Champion).
- Manesse, D. & Cogis, D. (2007): *Orthographe, à qui la faute?* Paris (ESF).
- Mortamet, C. & Gonac'h, J., (2011): Variation orthographique en français: le cas des non-natifs. In: Schnaffer, I. & Bertrand, O. (dirs.): *Variétés, variations et formes du français*. Paris (Les éditions de l'école Polytechnique).
- Mortamet, C. (2003): *La diversité à l'université, Analyse sociolinguistique de copies et de discours d'étudiants entrant à la faculté de Lettres et Sciences humaines de Rouen*. Thèse de doctorat, Université de Rouen.
- Porquier, R. & Py, B. (2005): *Apprentissage d'une langue étrangère: contexte et discours*. Didier (Paris).
- Selinker, L. (1994): *Reconsidering interlanguage*. London/New York (Longman).
- Tarone, E. (2005): Fossilization, social language and language play. In: Han, Z. & Odlin, T. (eds.): *Studies of fossilization in second language acquisition*, Clevedon, Multilingual Matters, 157-172.

De l'incidence du cotexte sur les choix orthographiques en FLE: étude de cas

Jean-Marc DEFAYS & Frédéric SAENEN

Université de Liège, Institut Supérieur des Langues Vivantes

This article is based on a survey carried out at the ISLV (University of Liège) among students in French as a Foreign Language (FLE) with a B1-level. The given exercise consisted of writing a sequence of words out of context and then of relocating the exact same words in the text from which they were taken. By means of a keen statistical analysis of the transcriptions, we attempt to determine where the most common spelling mistakes (lexical as well as grammatical) are made for a student in FLE, thereby assessing the impact of context on spelling.

Introduction

Il faut admettre que les recherches relatives à l'acquisition de l'orthographe, qui conjuguent et renforcent mutuellement les perspectives de la linguistique, de la psycholinguistique, de la linguistique cognitive, n'en sont encore qu'à leurs débuts, et qu'elles n'ont encore guère profité concrètement jusqu'à présent à la didactique des langues, maternelle ou étrangères. Il n'empêche que leur multiplication, leur diversification et leurs résultats, de plus en plus riches et intéressants, laissent penser qu'il ne tardera plus longtemps avant que les méthodes et les pratiques d'enseignement des langues puissent petit à petit se débarrasser de représentations, d'explications ou d'exercices inutiles voire néfastes au profit de stratégies pertinentes et efficaces.

C'est avec cette conviction que nous menons nos enquêtes sur l'acquisition de l'orthographe chez des apprenants de français langue étrangère, et que nous présentons ici un aperçu des indices sans doute encourageants qu'on peut en tirer, mais surtout des questions fondamentales que ces sondages, aussi limités soient-ils, soulèvent. Une des premières porte sur les rapports entre l'acquisition de l'orthographe en langue maternelle et en langue étrangère, soit dans l'absolu, c'est-à-dire sans tenir compte de la différence entre les systèmes des deux langues en présence (éventuellement plus de deux si l'on tient compte de langues-relais), soit en en tenant compte, par exemple quant à leur degré d'opacité relative. À ce titre on pourrait réactualiser les comparaisons déjà éprouvées entre l'apprentissage du code écrit de la langue maternelle et celui d'une langue étrangère, ou inversement entre l'apprentissage d'une langue étrangère et

les problèmes que rencontre le dysphasique ou l'illettré en langue maternelle.

La première démarche dans ce champ d'investigation consiste à dresser et revoir sans cesse la liste de facteurs qui peuvent influencer l'acquisition de l'orthographe. Parmi les principaux, on retiendra: le profil de l'apprenant (âge, scolarité, motivation,...); le niveau de maîtrise de la langue maternelle et de la langue étrangère cible qu'il a atteint; le système linguistique de la langue ou des langues qu'il maîtrise déjà et avec la(les)quelle(s) des transferts, conscients ou non, auront probablement lieu; le contexte situationnel, pragmatique de la performance; le cotexte syntagmatique, discursif, générique du ou des mots à orthographier; l'enseignement et les instructions qu'il a reçus concernant l'orthographe d'une manière générale ou concernant tel type de mots en particulier; etc. C'est la nature de ces facteurs, leur combinaison, leur évolution, et leur impact qui doivent faire l'objet de recherches minutieuses et méthodiques, telles que celle que nous avons tenté de mener ici et qui questionnent plus précisément les paramètres "cotexte", "langue-source", "orthographe d'usage vs grammatical" et "système français".

Fayol & Jaffré (2008) ont en effet bien mis en évidence qu'en matière d'apprentissage de la langue maternelle "peu de données fiables sont disponibles sur l'amélioration de l'orthographe lexicale par un enseignement explicite et systématique"¹. La remarque vaut également pour l'enseignement du FLE! La question que nous soulevons à travers les modestes résultats de notre étude de cas se situe à l'intersection du faisceau de difficultés au centre duquel se tient l'étudiant en FLE quand il s'agit d'orthographier un mot: influence de la langue maternelle / opacité de la langue cible, conscience de la coprésence de l'orthographe grammaticale et de l'orthographe d'usage, influence du cotexte sur la morphologie du mot, etc. Certes, notre recherche porte sur les productions d'un groupe d'étudiants restreint (une vingtaine), et ne concerne qu'un seul niveau (le B1 en l'occurrence); mais c'est, à notre avis, en analysant dans le détail les options graphiques des apprenants que nous sommes au plus près du diagnostic concret de leur difficulté et du pronostic d'une éventuelle remédiation.

1. Les hypothèses

Nous allons tenter de mettre à l'épreuve, à travers les résultats de notre recherche, les trois hypothèses suivantes:

¹ Fayol & Jaffré (2008:214).

1.1 Il y a des "zones de risques orthographiques" propres au public des apprenants en FLE:

acquérir une bonne orthographe en français n'a rien d'une sinécure pour les étudiants allophones. Il leur faut prendre en effet conscience de la coexistence d'une orthographe d'usage (souvent arbitraire ou, ce qui revient presque au même pour des étudiants non linguistes, explicable par des raisons historiques) et d'une orthographe grammaticale (plus justifiable, mais compliquée par les "nombreuses exceptions à la règle" qui relèvent, elles de l'arbitrarité).

1.2 Le cotexte amène une plus-value en ce qui concerne l'orthographe grammaticale:

le positionnement sur un axe syntagmatique d'un "mot" doit forcément être facilité par les éléments qui le précèdent et qui le suivent: déterminants, pronoms personnels, adjectifs, prépositions, etc. Non seulement il règle la question du choix sur l'axe paradigmatique (par exemple pour le cas des homophones) et dans la dimension sémantique, mais il éclaire également les flexions morphologiques ou les variations en genre et en nombre susceptibles d'affecter le terme. De nombreuses erreurs sont imputables au manque d'attention ou de compréhension que les étudiants manifestent face au cotexte, soit immédiat soit plus large, dans lequel s'inscrit l'unité de sens.

1.3 Les influences de la langue maternelle mais aussi de la /des langue-relais

sont prépondérantes sur les choix orthographiques, et c'est surtout le cas en ce qui concerne l'orthographe d'usage puisque, pour transcrire un mot qu'il n'aura jamais rencontré auparavant, un étudiant fera comme par réflexe le lien avec un autre terme qu'il connaît dans son propre idiome voire dans la *lingua franca* actuelle, l'anglais.

2. La démarche

Le présent article se base sur une recherche menée auprès d'une classe de 20 étudiants Erasmus de niveau B1, inscrits au cours du soir de FLE dispensé par l'*Institut Supérieur des Langues Vivantes*, à l'Université de Liège. Le public concerné se composait de neuf Espagnols (de E1 à E9), trois Italiens (de I1 à I3), deux Portugais (P1 et P2), un Anglais (An), un Allemand (All), un Bulgare (Bu), un Grec (Gr), un Chinois (Ch) et un Vietnamiens (Vi).

Le travail a été mené en deux temps. Dans une première phase (phase 1), les étudiants se sont vu dicter une liste de 45 mots de toutes natures, sans aucun environnement linguistique susceptible de les orienter vers le choix d'une orthographe. Il s'agissait de reporter sur une grille la première graphie qui s'imposait à l'esprit à l'écoute du vocable. Chaque mot a été dicté deux fois. La liste a ensuite été retirée aux étudiants. Cette phase a permis de mettre en évidence la première représentation graphique d'un mot, hors contexte, qui s'impose spontanément à l'esprit des étudiants testés.

Quant à la seconde phase (phase 2), un texte lacunaire a été distribué.² De ce document ont été retirés les mots dictés lors de la phase 1. Ils apparaissaient dans le document suivant le même ordre qu'en phase 1. Les étudiants ont dû remplir les blancs, en tenant compte cette fois du cotexte phrastique où apparaissaient les mots, et choisir en conséquence la forme qui leur paraissait correcte. Chaque paragraphe a été lu à haute voix une fois. On a redicté ensuite phrase à phrase. Enfin, on a procédé à une lecture à haute voix globale du texte, ininterrompue. Les étudiants ont donc entendu trois fois la forme manquante. Cette phase 2 a amené une dimension supplémentaire à l'analyse en ajoutant du cotexte, phrastique et textuel, puisqu'il s'agissait cette fois de réinsérer les vocables précédemment entendus dans un texte cohérent.

3. Les résultats

3.1 Phase 1

Le tableau suivant reproduit la liste intégrale³ des mots dictés aux étudiants. OT est l'Orthographe du Texte, telle qu'elle y apparaissait. AOA sont les Autres Formes Acceptées sous la plume de l'étudiant.

1. OT " cent " / AOA: sans, sens, sent, sang	24. OT " faisaient " / AOA: faisais/t
2. OT " réseau " / AOA: réseaux	25. OT " problématiques " / AOA: problématique
3. OT " ses " / AOA: c'est, s'est, sais, ces	26. OT " quelques " / AOA: quelque, quelqu'
4. OT " succès " / AOA: -	27. OT " déjà " / AOA: -
5. OT " point " / AOA: points	28. OT " télécharger " / AOA: téléchargé(e)(s)/ez
6. OT " créé " / AOA: créer, créés, créées	29. OT: " cas " / AOA: -

² Il s'agit de la première partie d'un article de presse, "Facebook pervertit-il les élèves?", signé Stéphanie Bocart et extrait du journal *La Libre Belgique* en ligne du 17 janvier 2011. Ce texte a été choisi car, selon son niveau de difficulté, son lexique, sa syntaxe et son thème, il appartient au corpus des documents susceptibles d'être proposés au cours de l'année par l'enseignant à ses étudiants de niveau B1.

³ Afin de respecter les contraintes de longueur de cet article, seuls les résultats les plus significatifs qui sont exploités dans nos conclusions ont été conservés.

7. OT "Belges" / AOA: belge(s), Belge	30. OT "illégal" / AOA: illégale/s
8. OT "quatre" / AOA: -	31. OT "lèvent" / AOA: lève(s)
9. OT "scolaire" / AOA: scolaires	32. OT "âge" / AOA: âges
10. OT "jeux" / AOA: jeu, je	33. OT "mis" / AOA: mi
11. OT "innombrables" / AOA: innombrable	34. OT "technologie" / AOA: technologies
12. OT "règles" / AOA: règle	35. OT "sein" / AOA: saint/s, sain/s, cinq
13. OT "bancs" / AOA: banc	36. OT "où" / AOA: ou
14. OT "installés" / AOA: installer, installé(e/s)	37. OT "tellement" / AOA: -
15. OT "pris" / AOA: prix, prit	38. OT "outils" / AOA: outil
16. OT "reçoivent" / AOA: reçoive(s)	39. OT "permettent" / AOA: permette/s
17. OT "visite" / AOA: visites	40. OT "enquête" / AOA: enquêtes
18. OT "également" AOA:-	41. OT "possèdent" / AOA: possède/s
19. OT "tant" / AOA: t'en, temps, tend(s)	42. OT "dont" / AOA: don
20. OT "dû" / AOA: du, du(e)(s)	43. OT "accès" / AOA: axais/t/ent
21. OT "régler" / AOA: réglé(e)(s), réglez	44. OT "contrôle" / AOA: contrôles
22. OT "problèmes" / AOA: problème	45. OT "fréquents" / AOA: fréquent
23. OT "commençaient" / AOA: commençais/t	

On remarque d'emblée la disparité des termes dictés, tant au niveau de leurs natures que des difficultés orthographiques qu'ils posent.

Cette liste comporte en effet:

- des termes dont **une seule orthographe** est **admissible**: par exemple "succès" ou les adverbes;
- des substantifs, adjectifs ou déterminants qui peuvent présenter **deux ou trois orthographes, selon la variation en genre et/ou en nombre**: "réseau", "règles", "bancs", etc.
- des **formes verbales** dont l'orthographe change selon une **variation de personne**: "reçoive/es/ent", "commençais/ait/aient", etc.
- des **formes verbales** dont les variantes orthographiques relèvent d'une **variation modale** (infinitif, participe passé, certaines formes de l'indicatif): "installés", "régler", etc.
- des termes dont l'**homophonie** peut engendrer des graphies concernant **deux natures différentes**: belge / Belge, pris / prix, dû / du, etc.
- des termes qui ont une **homophonie très riche**, car transversale à de **nombreuses natures**: "cent", "ses", etc.

Pour nombre de ces formes se posent de surcroît les problèmes traditionnellement liés à l'orthographe d'usage: présence de double consonne? accents? digrammes / trigrammes? nasales? hiatus? lettres muettes internes ou finales?

3.2 *Graphies proposées par les étudiants en phase 1, puis en phase 2*

Pour chaque famille de mots définie ci-dessus, nous allons présenter et analyser les performances des étudiants en sous-catégorisant tout d'abord les problèmes ponctuels liés à la perception du rapport oral / écrit ainsi qu'à l'orthographe d'usage (phase 1) puis en révélant l'incidence de la mise en cotexte sur leurs graphies définitivement adoptées (phase 2). Par souci de commodité de lecture et de clarté, les nombres sont écrits en chiffres et non en lettres. Les identifiants d'étudiants (E1, E2, etc.) sont en général directement suivis de leur graphie entre guillemets. Le descriptif de la phase 2 est introduit par la flèche et par la transcription du cotexte phrastique. Les initiales PC signifient "Proposition correcte".

3.2.1 Les termes dont une seule orthographe est admissible

Le substantif "**succès**" n'a d'emblée été orthographié correctement que par 7 étudiants. E9, P1 et All orthographient à l'anglaise "success". Le problème principal est lié à l'accent grave du mot et au redoublement de la consonne "c".

⇒ "Et le **succès** auprès des jeunes...": **6 PC**, de la part des étudiants qui avaient proposé "succès" en phase 1. E3 et Ch, qui n'avaient rien perçu en phase 1, proposent cette fois "success". Aucune des graphies fautives relevées en phase 1 n'est corrigée, parfois même elles sont aggravées (E7 passe de "sucedde" à "sucsede", I1 de "suxes" à "suxece").

Le substantif "**cas**" a semble-t-il beaucoup dérouté les apprenants. Si 10 graphies sont correctes, les 10 autres dénotent un embarras: sur les 5 Espagnols qui ont hésité, E6 et E8 optent pour "quand". An et All transcrivent "car" (conjonction ou anglicisme?) et E4 l'anglicisme "calm". P1 écrit phonétiquement "ca" et E9 "ka".

⇒ "...dans la plupart des **cas**,... ": **15 PC**. 7 étudiants rectifient le cap, et parfois pour des erreurs lourdes ou des absences de réponses en phase 1. Plus gênants sont les problèmes de E1 et de E3: E1 avait proposé "cas" en phase 1 mais il modifie en "càs"; E3 avait lui aussi identifié le mot en phase 1, mais il n'intègre rien à cet endroit du texte lacunaire.

Le numéral "**quatre**" n'a été mal orthographié que par I2 "quattro", par interférence.

⇒ "... et **quatre** autres ados...": **19 PC**. Idem phase 1.

L'adverbe "**déjà**" a été correctement orthographié par 14 étudiants. Les problèmes rencontrés par les six autres concernent bien sûr la succession accent aigu / accent grave.

⇒ "... sont **déjà** identifiés...": **15 PC**. Les 6 étudiants pointés en phase 1 commettent à nouveau une erreur. Tous (sauf E3 et P2 qui maintiennent leur "dèja" initial) tentent de rectifier, mais ils n'y parviennent pas. I1 passe de "déjà" à "dejá", I2 de "dejà" à "dejá", I3 de "dêja" à "dèja" et P1 de "dejá" à "déjá").

Les **adverbes en "-ment"** sont intéressants à observer. "**Tellement**" est correctement écrit 12 fois. Il est proposé avec des accents chez 3 hispanophones (E1: "tellèmènt", E6: "tèlement"; E7: "téllement"). Aucun italophone ne semble avoir appliqué la règle de formation sur l'adjectif féminin (I1 et I3 "telment", I2 "talment"). P1 "télemont" propose une orthographe qui ne montre aucun lien avec la forme adverbiale. "**Également**" est correctement écrit 11 fois. Il ne subit qu'une interférence, chez E7 "également". I1 "égalmant" perd le lien avec l'adverbe, qu'il avait pourtant maintenu en finale dans "telment". Ni P1 ni P2 n'appliquent la règle de formation et donnent "égalment". Parmi les étudiants hors du domaine roman, seul All propose une forme erronée, avec "égalément".

⇒ "...ce n'est plus **tellement** Internet...": **12 PC**. Les étudiants qui avaient proposé "tellement" en phase 1 réitèrent la performance. Seul E1 corrige parfaitement son accentuation initialement fautive, ce qui amène le total à 13 PC. Les erreurs relèvent d'accents ajoutés ou déplacés à mauvais escient (E8 aboutit à "tèlement") ou de formation partielle (I1 "talement" féminise la forme italienne de l'adjectif!). E6 maintient "tèlement" et I2 "talment". La rectification que P1 propose à "télemont" en "télèmant" s'éloigne encore un peu plus d'une identification avec la règle de formation des adverbes en "-ment". Enfin, E9 rectifie son orthographe correcte de phase 1 en "téllement".

⇒ "... qu'il adresse **également** aux enseignants...": **12 PC**. Seul E9 corrige avec fruit son "égalment" de la phase 1. P1 et P2 maintiennent par contre cette graphie et I3 l'adopte pour préciser quelque peu son "eglement" de phase 1. E7 conserve "également". Seul All redouble le "l" dans sa nouvelle proposition, avec "egallement". I1 perd le lien avec la base adjectivale "égal" en passant à "eguelmant". Enfin, deux étudiants qui avaient en phase 1 une bonne graphie introduisent une erreur: E1 "égalèment" et I2 "egalement".

3.2.2 Les substantifs, adjectifs ou déterminants qui peuvent présenter deux ou trois orthographes, selon la variation en genre et/ou en nombre

Dans cette catégorie, en phase 1, certains termes n'ont pour ainsi dire posé aucun problème: c'est le cas de "**visite**" (faux seulement chez E4 "visit"), de "**point**" (partout correct sauf I3 "pain" et P1 "puant") et de "**quelques**" (faux chez seulement Ch "qualeque"). Un mot comme "**problèmes**" sur lequel il

avait été insisté en classe lors de cours précédents est aussi globalement réussi (correct dans 17 cas).

⇒ "L'objet de sa **visite**...": **20 PC**. Le mot est partout correct, même chez E4 qui le rectifie.

⇒ "... à tel **point** qu'aujourd'hui près de 100% des ados...": **19 PC**. P1 modifie "puant" en "pont".

⇒ "... **quelques** points noirs...": **10 PC**. Ch n'identifie toujours pas le mot et écrit "quel". 9 étudiants (E3, E4, E5, E7, E9, I2, P1, P2 et All) omettent de le mettre au pluriel, alors que le cotexte est clairement pluriel.

⇒ "... à cause de **problèmes** liés...": **10 PC**. L'évidence n'est pas apparue aux étudiants dans ce cas en raison du "de", souvent envisagé comme singulier, pourtant 11 étudiants l'indiquent. Les autres maintiennent leur orthographe de départ, correcte ou erronée.

a) Les substantifs

"**Banc**" est le substantif qui a été globalement le moins bien identifié. Seul E1 en propose une graphie acceptable. 16 étudiants n'identifient pas la nasale. Ainsi, tous les I et les P, 5 Espagnols et An proposent "bon". Parmi les solutions les plus éloignées, on trouve Gr "bot", Ch "bomb" et E9 "bine". Bu et Vi proposent l'anglicisme "band", All propose "ban"⁴. E3 et E4 sont les plus éloignés car ils donnent "vont".

⇒ "... trois longs **bancs** de bois...": **3 PC**. E1 maintient la forme correcte de phase 1 et on observe quatre corrections qui aboutissent à la graphie attendue (E2, I2) soit à l'identification du mot "banc", mais au singulier (E7, I3). Ces quatre étudiants corrigent l'identification de la nasale, ce qui est aussi le cas de E5 "bandes", E8 "banques", I1 "bande", Gr "band", bien qu'ils aboutissent à des mots inappropriés. E3, E6 et All n'indiquent rien cette fois. Ch et Vi maintiennent des formes erronées, mais les accordent au pluriel. P2 "bôll" aboutit à la forme la plus aberrante.

"**Outil**", du fait de la diphtongue et de la lettre finale muette, a aussi perturbé les étudiants. Seuls 8 en connaissaient manifestement l'orthographe correcte. 5 proposent "outi". E8 et I3 transcrivent "uti(s)". À noter qu'aucun étudiant n'a, à ce stade, commis la confusion fréquente avec l'adjectif "utile".

⇒ "...la mise à disposition pour des prix tout à fait raisonnables d'**outils** mobiles...": **7 PC**. 7 étudiants qui avaient en phase 1 correctement orthographié le mot aboutissent à la forme correcte au pluriel requise par le cotexte. E6 corrige "outis" en "outils". La confusion avec l'adjectif "utile"

⁴ Mais c'est une option phonétique car, en tant que B1, il ne connaît sans doute pas l'existence de ce terme rare en français...

apparaît dans 5 cas chez des étudiants qui, en phase 1, avaient proposé "outi" (E3), "outir" (I1), "outi" (P2), rien (E5 et P1). E9 et All combinent les deux formes en "outile". La diphtongue "ou" pose des problèmes évidents de transcription, dans la mesure où I2 passe de "outi" à "otils" et P1, qui n'avait rien proposé, donne "úti".

"**Contrôle**" a de même été correctement transcrit par 5 étudiants. La majorité des autres ont subi l'interférence de l'anglais, pour la perte du circonflexe (maintenu chez E4 et P1 "contrôl" et All "contrôle") et du e final.

⇒ "... une perte de **contrôle**...": **4 PC**. Il s'agit de An, Bu et Vi (qui avaient déjà prouvé en phase 1 maîtriser l'orthographe du terme) et de Ch qui corrige "control". Tous les autres sont tributaires de l'anglicisme ou de leur forme erronée de phase 1. Deux étudiants se corrigent à mauvais escient: E1, qui passe de la forme correcte en phase 1 à l'anglicisme "control", et Gr, qui propose "controle". E7 fait apparaître le circonflexe sur la syllabe nasale.

b) Les adjectifs

"**Innombrable**" a été correctement orthographié par tous les E, sauf E6. Les difficultés des autres étudiants n'ont pas nécessairement porté sur le redoublement (présent chez 4 fautifs sur 8) mais bien sur l'identification de la nasale (on obtient ainsi I3 "innobtable", All "ineaubrable"), voire sur la syllabe finale (Vi "innonrable", I1 "innomabre").

⇒ "... les services **innombrables**...": **6 PC**. Sur les 10 étudiants qui avaient correctement écrit ce mot en phase 1, seuls 5 pensent à l'accorder au pluriel, malgré le cotexte évident. I1, qui partait de "innomabre", rectifie non seulement l'orthographe mais l'accorde correctement. En général, les étudiants qui avaient en phase 1 une orthographe incorrecte la maintiennent. All affine son "ineaubrable" en "inonbrable". Par contre le cotexte ne semble pas aider An qui passe de "innoblable" à "enomblable". Ch, qui n'avait rien proposé initialement, donne "inombrales".

Alors que, rappelons-le, "problème" n'est mal orthographié que 3 fois sur 20, "**problématique**" est mal orthographié 13 fois sur 20! Ce renversement s'explique par l'oubli de l'accent (8 fois) ou son maintien en accent grave (5 fois). S'ajoute la question de la finale "-ique", que E8 et E9 anglicisent en "-ic".

⇒ "... une série de **problématiques** qui ne pouvaient...": **4 PC**. Le substantif est lui aussi précédé d'un "de" pluriel mais seul le verbe, un peu plus loin, porte un indice de nombre. Dès lors, l'accord n'est effectué que par 5 étudiants ici! Les erreurs décrites en phase 1 (accents mal orientés ou omis, finales en -ic) se maintiennent.

"**Illégal**" posait le double problème de l'accent et du redoublement consonantique. 8 étudiants seulement proposent une des formes admises. Quand les étudiants mettent les "ll" (E1, I2, All, Bu, Gr), ils se trompent avec l'accent, soit omis ou orienté en grave. E5 donne "ilègal". La graphie de Ch "éligal" est la plus éloignée et trahit une difficulté de perception phonétique.

⇒ "... le téléchargement est totalement **illégal**...": **2 PC**. À l'exception de Bu, aucun étudiant n'arrive à une orthographe correcte de ce mot en cotexte. 5 l'accordent au féminin, All va jusqu'au féminin pluriel "illéga**l**es". Chez E3, e7 et I3, l'accent aigu indiqué en phase 1 disparaît. E9 réduit le double "l" à un seul. La plupart des orthographes incorrectes de phase 1 sont maintenues, seuls E5 et Gr corrigent leur accent grave en accent aigu.

"**Fréquent**" n'a été correctement orthographié que 7 fois. Pour E3, E5, E6, I3 et All, l'accent aigu est absent. Bu propose une mauvaise terminaison, "fréquant". E7, I1 et I2 donnent la forme au féminin, ce qui ne correspond pas à la réalité phonologique dictée. E4 et Vi confondent avec "fréquence".

⇒ "Les dérapages sont de plus en plus **fréquents**." **4 PC**. Bu corrige "fréquant" en "fréquents". Le mot "dérapages" est apparemment perçu comme féminin par 7 étudiants, qui accordent l'adjectif en fonction. Parmi ceux-là, on retrouve E4 et Vi, qui partaient du substantif "fréquence" et le corrigent en adjectif. Les problèmes d'accents sont corrigés chez E3, I2, I3 et P2. Par contre, l'accent disparaît chez E7.

3.2.3 Les formes verbales dont l'orthographe change selon une variation de personne

La forme "**reçoivent**" (acceptée au singulier) n'a pas posé de difficulté à 9 étudiants. Les graphies erronées s'échelonnent comme suit: la cédille oubliée par Gr "recoivent"; une forme hybride avec l'idée d'infinitif chez E7 "reçoivre" ainsi que E9 et All "recoivre"; des résolutions graphiques du "ç" et de la diphtongue "oi" en E4 "résuave", E6 "resuave", P1 "ressouave", P2 "reçuavre"; un calque approximatif de la langue maternelle pour I1 "recevre"; le retour à l'infinitif chez I2 "récevoir" et Ch "recevoir".⁵

⇒ "Ce matin, ils **reçoivent** Christophe Bustraen...": **11 PC**. E1, qui avait proposé une forme correcte en phase 1, se corrige mal en écrivant ici "réçoivent". Les deux corrections les plus saillantes sont E6 et E7. I2 et All donnent "recoivent", prouvant qu'ils ont intégré le fait que le verbe devait être conjugué. E9 écrit cette fois "reçois". Les deux orthographes les plus

⁵ Signalons que, dès les premières leçons dispensées au B1, la conjugaison du présent de l'indicatif est revue en profondeur et censée maîtrisée pour la composition des autres conjugaisons (imparfait, subjonctif...).

éloignées de la forme correcte et empreintes de phonétisme pur sont à nouveau P1 "résouvae" et P2 "reçave".

La forme "**commençaient**" est la seule à avoir été ratée par tous les étudiants. En effet, pas un seul n'a identifié un imparfait lors de la dictée. 10 ont proposé l'infinitif correct et 1 l'infinitif incorrect (I2 "commancer"). 6 autres ont proposé le participe passé, au masculin, correct, et 1 incorrect (P1 "commancé"). I2 écrit "comonce".

⇒ "...trois adolescents qui à cette époque-là **commençaient** à surfer...": **4 PC**. Ce sont E1, E2, E8 et Bu. Gr propose "commençait". E3, E4, E5, E7 proposent le présent "commencent" et I2 "commacent". E6, E9, P2, An y voient le participe passé, tout comme I1 "comoncé" et I3 "commoncé" qui commettent des erreurs d'usage. All, Ch et Vi maintiennent l'infinitif.

Pour "**faisaient**", aucun étudiant n'a pensé à la graphie du pluriel (la plus ardue, en général) mais 8 ont proposé des formes acceptables ("faisais" ou "faisait"). E1 a redoublé le "s". P1 et P2 ont proposé une conjugaison erronée du présent "faisez" E3, E5, I1 et CH n'ont rien écrit. Les autres orthographes conservent l'idée de forme verbale, conjuguée ou à l'infinitif, mais sont assez farfelues (E8, "faissez", I3 "faisév", E9 "faussait", E2 "fosser", All "feuser").

⇒ "C'est en m'intéressant à ce qu'ils **faisaient**...": **6 PC**, de la part des étudiants qui avaient proposé des formes acceptables en phase 1. E4 et E7 passent de "faisait" à "faisent". E1 maintient le "s" redoublé et accorde au pluriel. P1 et P2 ajustent en P1 "fasait" et P2 "faisent", se rapprochant du sémantisme du verbe. Les autres orthographes farfelues se modifient également, mais jamais pour arriver à la correction parfaite: E9 donne "faissaient", E2 et All "faisait", I3 "faisent", E8 "faissent"). E3, qui n'avait rien en phase 1, propose cette fois "faisant". Ch ne propose toujours rien.

3.2.4 Les formes verbales dont les variantes orthographiques relèvent d'une variation modale

"**Créé**" a d'abord été perçu comme l'infinitif (11 propositions correctement orthographiées; E1 "créér" et E3 "crèer"). E8 propose l'alternance "créer / créé". Les problèmes relèvent de la position de l'accent (All "créé" et I1 "crée"), du hiatus résolu (I3 "cré"), de la confusion avec un autre verbe l'infinitif ou conjugué (Ch: "crier" et P1 "creiez").

⇒ "...des adolescents ont **créé** leur profil...": **8 PC**. An maintient sa proposition "créé" de phase 1 et 7 étudiants, qui avaient tous proposé l'infinitif correct en phase 1, aboutissent au participe "créé". En dehors de cela, les tendances sont très diverses. L'accentuation du participe a posé le principal problème: E1, E2 et All transcrivent "crée", E3 donne "créés" et E9 "créé". Le hiatus pose un problème à E5 qui, bien qu'ayant proposé "créer" en phase 1, modifie en "creié". E4 "créent" et I1 "creient" proposent

une forme logiquement incompatible avec l'auxiliaire conjugué déjà présent dans la phrase. Ch corrige en "écrié" et P2 passe de "créer" à "criée".

3.2.5 Les termes dont l'homophonie peut engendrer des graphies concernant deux natures différentes

Dans le mot "**Belges**"⁶, seule la majuscule distingue la nature. E2, E8, E9, P1, An, Bu et Ch ont commencé le mot par la majuscule. On notera E3 "belgue" et E8 "Bèlge".

⇒ "Comme des milliers de jeunes **Belges**...": **2 PC**. Seuls P1, An et Bu marquent cette fois la majuscule qui distingue le substantif de l'adjectif. Le pluriel est respecté dans 16 cas, seuls l'oublie P1, P2 et Ch. E3 "belgues" reproduit l'erreur de la phase 1.

"**Pris**", représenté 8 fois, a été supplanté par son homophone nominal "prix" (12 fois).

⇒ "Les sept ados y ont **pris** place": **16 PC**, bien qu'il figure dans une expression figée ("prendre place"). I1 maintient "prix", P2 passe de "prix" à "prit" et Ch de "prix" à "prie".

"**Dû**" a été adopté a priori par 9 étudiants, 11 autres ont opté pour la contraction "du". Seuls E1 "doux" et Ch "tu" ont commis des confusions phonétiques.

⇒ "...les deux premières situations que j'ai **dû** régler...": **9 PC**. Parmi elles, 6 étudiants ont articulé avec l'infinitif "régler". E3, E5, E9, I2 et P1 maintiennent "du". I1 et All passent de "dû" en phase 1 à "du" ici. Ch et E1 corrigent. E7, qui avait proposé "dû" en phase 1, ne propose rien ici.

"**Mis**", choisi a priori par 16 étudiants.

⇒ "Moi, j'ai **mis** que j'avais 2 ans...": **18 PC**, sauf All qui passe de «mis» en phase 1 à "mi".

Pour "**où**", étonnamment, c'est la graphie du pronom relatif "où" qui a été choisie en majorité (14 occurrences!) avant celle, plus simple puisque dépourvue d'accent, de la conjonction de coordination "ou" (4 occurrences). I3 et P1 proposent erronément "où".

⇒ "le service tout en un de Facebook **où** il y a possibilité...": **15 PC**. I3 et P1 fossilisent "où". I1 et Ch maintiennent "ou".

"**Dont**", a été correctement proposé 9 fois. E1, E2, E6, E7, I3, All confondent avec "donc". Ch interprète correctement la nasale mais pas la consonne initiale et propose "ton". Les graphies les plus aberrantes sont P1 "dan", I1 "dôn" et E8 "d'eau".

⁶ Nos étudiants sont forcément familiarisés avec ce vocable!

⇒ "**dont** 63 % disposent d'un accès": **8 PC**. 6 étudiants qui avaient déjà proposé "dont" en phase 1 reproduisent cette réponse. Seuls E8 "d'eau" et P2 "don" corrigent en "dont". Les autres maintiennent "donc", ou le proposent cette fois, comme I1 ou Ch.

Enfin, le substantif "**accès**" a été classé ici, car il offrait une homophonie avec certaines formes de l'imparfait de "axer". Aucun étudiant n'a opté pour l'une de ces dernières, mais le x de ce verbe, ou de son participe passé, était présent chez E2 "axé", I1 "axes", I2 "axces". Les autres erreurs portent sur l'absence d'accent ou la confusion avec l'anglais (E9 et All "access"). An a compris "excès".

⇒ "...disposent d'un **accès** Internet...": **6 PC**. L'ambiguïté avec le verbe "axer" ou le mot "excès" est levée pour E2 et An (qui corrigent) et I1 "acesse", mais pas pour I2 "axcès". Le problème de l'accent final, absent dans la plupart des cas ou mal orienté chez E4, se pose pour 11 étudiants. Le mot anglais "access" est choisi par E9 et All.

3.2.6. Les termes qui ont une homophonie très riche, car transversale à de nombreuses natures

"**Cent**" a suscité 16 graphies correctes. E9 et Bu donnent d'emblée la graphie attendue dans le texte, sinon, le choix graphique le plus répandu est "sans", pour 6 étudiants. A l'inverse de la tendance observée avec "banc", la nasale a été parfaitement identifiée sauf dans trois cas: I1 et I3 "son" et P1 "sont". E8 marque une hésitation entre "sans" et "sain". E4 propose le mot anglais "sound".⁷

⇒ "Cinq **cent** millions. " **16 PC**. E7, E9, Bu et Ch accordent en "cents".

Les formes suscitées par "**ses**" sont globalement acceptables, même si aucun étudiant n'a a priori pensé à la forme du déterminant possessif. On trouve le démonstratif "ces", les formes conjuguées "sais / s", les combinaisons présentatives "c'est" et réflexive "s'est". I1 donne "cet" et E9 "ce".

⇒ "...discuter en ligne avec **ses** amis...": **17 PC**. Seuls E5, E7 et E8 maintiennent "ces".

"**Sein**" a posé de réelles difficultés. Seul Gr a proposé la graphie du texte. Les autres formes admises étaient Bu "saint", Vi "sain", E2 E4 et All "cinq". E1, E7, I3 et P2 donnent "sens". E8, E9 et I1 donnent "sans". P1 donne

⁷ A noter qu'une incertitude règne quant au choix de la graphie "sens" car E2, E5 et I2 l'ont peut-être choisie en pensant au substantif "sens" et non à la forme conjuguée de "sentir"...

"cent" et An "sang". E3 propose la seule graphie dénasalisée avec "scène" et Ch la forme inexistante "sin".⁸

⇒ "au **sein** de la police": 5 PC. L'expression ne permet pas de lever l'ambiguïté ou d'éclairer le sens. 4 bonnes réponses seulement, dont Gr. E2 et All passent à "seine". 7 étudiants proposent "sens" ou "sense", soit qu'ils le maintiennent, soit qu'ils l'avancent après s'être abstenus en phase 1 (comme E5 et E6). Les autres réponses se font en dépit du sens contextuel ou de la phonétique : E9 "chef", I1 "centre", P1 "sede", An "s'en". E3 s'abstient cette fois de répondre.

Enfin, "**tant**" n'a été spontanément transcrit que par Gr et Vi. Parmi les propositions acceptables, on compte 8 étudiants qui ont opté pour "temps". S'il n'a jamais été confondu avec "dans", le mot a suscité bon nombre de graphies incorrectes. E5, E7, E9 et P1 donnent "temp". E2 et All inventent "tent". I2 propose "tot" et E6 est le seul à confondre la nasale avec "ton".

⇒ "En **tant** que médiateur...": 14 PC, dont 12 corrigées de "temps" ou de formes éloignées. Seuls E5, E7, E8, A, et Ch maintiennent ou proposent "temps".

4. Bilan de la recherche par rapport aux hypothèses de départ

Le tableau synoptique ci-dessous, élaboré à partir des résultats obtenus en phase 2, propose un regroupement de certains mots de l'exercice selon deux critères, afin de nourrir la réflexion qui suit:

<i>Mots les mieux réussis:</i>	<i>Mots les moins bien réussis:</i>
visite (20 PC)	Belges (2 PC)
point (19 PC)	illégal (2PC)
quatre (19 PC)	bancs (3 PC)
mis (18 PC)	contrôle (4 PC)
pris (16 PC)	fréquents (4 PC)
âge (16 PC)	problématiques (4PC)
déjà (15 PC)	commençaient (4 PC)
où (15 PC)	sein (5 PC)
	accès (6 PC)
	succès (6 PC)
	faisaient (6 PC)

Les constats précédents, confrontés aux hypothèses de départ de notre étude, nous permettent d'avancer les conclusions suivantes:

⁸ Comparer les choix pour "cent" et "sein" permet de mesurer le désarroi des étudiants face aux homophones courts, car certains qui avaient bien identifiés le premier en appliquent une graphie équivalente au second.

4.1 *Il y a des "zones de risques orthographiques" propres au public des apprenants en FLE*

Il est clair que les problèmes d'orthographe d'usage les plus fréquemment rencontrés par les étudiants FLE sont relatifs à la non biunivocité dans le rapport oral/écrit du français. On y retrouve ainsi les doubles consonnes, les digrammes ou trigrammes censés rendre un seul son, la transcription des voyelles nasales, les lettres muettes, les accents et signes diacritiques, etc.

La conjonction d'exigences morphologiques et flexionnelles dans un même mot dès qu'il est situé en contexte induit souvent l'étudiant en erreur, ce qui explique que les formes verbales comme "commençaient" par exemple aient été si mal réussies: elle cumule un redoublement, une cédille, la flexion du pluriel en -nt et la nécessité de discriminer phonologiquement la finale par rapport à l'infinitif ou au participe passé.

Parmi les mots les moins bien réussis, on retrouve également "banc", qui a suscité la perplexité dès la phase 1 et n'a guère été éclairci par le cotexte. Il faut dire que les complexités y foisonnent: le /b/ explosif initial est souvent confondu avec /v/ par les hispanophones, la nasale est perçue différemment aussi bien phonétiquement que dans son rendu orthographique, la dernière lettre est muette, le signifié n'est pas connu de tous les apprenants B1 et, cerise sur le gâteau, il faut penser à mettre le pluriel! Ce graphème très court est emblématique des "nœuds" de problèmes orthographiques que recouvrent certains mots, à différents niveaux, et qui déstabilisent complètement les apprenants.

4.2 *Le cotexte amène une plus-value en ce qui concerne l'orthographe*

Il est clair que, pour ce qui concerne l'orthographe d'usage, la mise en cotexte est relativement inopérante pour la rectification d'orthographe ignorées. Les erreurs fossilisées⁹ le demeurent, en dépit de l'environnement du mot.

La contextualisation peut par contre s'avérer fort utile pour discriminer le bon choix à opérer parmi des homophones sur un axe paradigmatique autant que syntagmatique. C'est le cas de "cent", "tant" et "ses" (pas du tout de "sein", terme utilisé dans une expression figée et mobilisant une nasale difficile à identifier pour les étudiants). Ces mots sont, avec "cas", les mieux corrigés dans la phase 2.

⁹ On entend par fossilisation lorsque les "imprécisions et [l]es incorrections [sont] incrustées au cours de l'apprentissage [...] parce qu'elles n'ont pas été corrigées à temps". D'après Defays, J.-M. & Deltour, S. (2003): *Le français langue étrangère et seconde*. Enseignement et apprentissage. Sprimont (Mardaga).

Les participes sont aussi très bien appliqués, dans la mesure où "mis" souffre peu de concurrence à un niveau B1 (les étudiants ne pensent pas d'emblée à la notation musicale ou à l'expression de la moitié) et où l'ambiguïté entre "prix" et "pris" est vite dissipée par le sens de la phrase. Dans le cas du mot "Belges", le choix de la nature nominale dénotée par la majuscule n'est pas facilité par le cotexte, car la proximité immédiate de l'adjectif "jeunes" a sans doute été contaminante...

On remarque cependant que, en dehors des environnements très clairs, surchargés en indices, la mise en cotexte ne conditionne pas systématiquement le choix dans ce qui constitue une clé de voûte de l'orthographe grammaticale: les marques du pluriel. L'article "de / d'" n'est ainsi pas nettement perçu comme introducteur d'un pluriel.

Les difficultés entraînées par les formes à l'imparfait "faisaient" (6 PC) et "commençaient" (4 PC) sont plus inquiétantes car elles montrent, par le maintien de formes connexes (infinitifs, participes...) voire aberrantes, la déconnexion qui peut persister entre le mot et le sens du texte. Pour "commençaient", la difficulté à choisir la conjugaison provenait sans doute en grande partie de l'écran entre le sujet et le verbe. L'indicateur temporel aurait cependant pu mettre sur la voie...

En fait, tout se passe comme si les apprenants ne parvenaient pas à embrayer entre ces deux vitesses que seraient l'orthographe grammaticale et l'orthographe d'usage, quand elles posent plusieurs problèmes simultanément dans le même mot.

4.3 Les influences de la langue maternelle mais aussi de la langue-relai sont prépondérantes sur les choix orthographiques

Debyser (1970) définit d'office, du point de vue de la pédagogie des langues vivantes, l'interférence comme une "faute". Or, les influences de la langue maternelle ou relai, inévitables, peuvent parfois jouer un rôle de facilitateur, donc positif (par exemple dans le cas de "point", mot identique à son équivalent anglais, maîtrisé par les hispanophones), même si les cas restent rares en français. On le voit avec "illégal", l'un des mots les moins réussis (2 PC) alors qu'il existe phonétiquement quasi à l'identique en italien ou en espagnol, mais les "ll" et l'accent perturbent ici le jeu.

Les interférences négatives se vérifient par exemple avec le mot "contrôle", correct dans 4 cas seulement, à cause de sa graphie anglicisée! Et "succès" n'est pas loin de subir le même sort, si ce n'est qu'il suscite des orthographes autrement hérissées de la part de ceux qui ne se basent pas sur l'anglais pour le former...

Dans le cas des adverbes en "-ment", on s'aperçoit que les graphies incorrectes se concentrent chez des locuteurs de langue romane,

notamment italophones. Plus subtilement encore, on relève le cas du déterminant indéfini "quelques", laissé au singulier malgré son environnement pluriel. Ce phénomène s'explique par l'interférence, en tout cas pour les langues romanes dans lesquelles la construction reste au singulier.

5. Conclusions et perspectives

Des travaux comme ceux qui précèdent sont révélateurs d'une multiplicité de facteurs d'explications au problème de la "cacographie" et de l'intrication des problèmes spécifiques que pose la langue française sur ce point. L'orthographe est souvent envisagée comme une fatalité par l'apprenant en FLE, qui hérite comme par contrecoup des discours stéréotypés selon lesquels "Le français, c'est une langue pleine d'exceptions", "En français, rien ne s'écrit comme ça se dit", etc.

Deux nécessités semblent s'imposer à l'enseignant face aux difficultés orthographiques de ses étudiants. La première est bien entendu la description, la comparaison et l'explication d'erreurs, chez différents apprenants de différentes langues dans différentes circonstances, de façon à définir et à redéfinir des typologies, et des situations ou des zones "à risques", qui pourront aider l'enseignant à (faire) comprendre les phénomènes et les mécanismes concernés par ces erreurs, et à les traiter, voire les prévenir. La seconde nécessité, plus utile encore, est celle de la mise à l'examen, à l'épreuve, à l'œuvre de stratégies, à la fois d'enseignement et d'(auto)apprentissage. On a encore beaucoup à découvrir sur les opérations de segmentation, d'appariement, d'assemblage auquel se livre, spontanément ou non, un enfant ou un adulte confronté à une langue étrangère; sur la constitution d'un lexique mental nouveau et de ses interférences ou transferts avec celui de la langue maternelle; sur la distinction entre la transcription phonologique et la récupération directe (adressage lexical), et sur leur succession, leur interaction, leur combinaison lors de l'apprentissage d'une langue étrangère.

Puisque le cœur du problème semble se situer notamment à l'articulation entre orthographe d'usage et orthographe grammaticale, voire dans le choix entre voie directe (lexicale) et indirecte (non lexicale)¹⁰, ne s'agirait-il pas de s'interroger de façon constructive – avec les étudiants et sur base de leurs difficultés à discriminer les sons d'un mot, à identifier sa nature hors et en contexte – afin de trouver une dynamique à acquérir entre ces deux dimensions qui sont vouées à cohabiter dans le langage écrit, ceci dans le cadre d'un apprentissage raisonné de l'autoanalyse et de

¹⁰ Voir à ce sujet l'article de Pacton, S., Fayol, M. & Perruchet, P (1999) mentionné en bibliographie.

l'autocorrection, à la charnière épilinguistique entre la réflexion métalinguistique et la pratique linguistique ? C'est à cet endroit qu'à notre avis les analyses qui précèdent encouragent à entreprendre des recherches plus poussées et des expériences riches d'enseignement.

Bibliographie

- Castelloti, V. (2005): La Langue maternelle en classe de langue étrangère. Paris (CLE International).
- Catach, N. (1995): L'orthographe française. Paris (Nathan).
- Cuq, J.-P. dir. (2003): Dictionnaire de didactique du français. Langue étrangère et seconde. Paris (CLE International).
- Debyser, F. (1970): La Linguistique contrastive et les interférences. In: Langue française, Volume 8, n°1, 31-61.
- Defays, J.-M., Maréchal, M. & Melon, S. (2000): La Maîtrise du français du secondaire au supérieur. Bruxelles (De Boeck-Duculot).
- Defays, J.-M. & Deltour, S. (2003): Le français langue étrangère et seconde. Enseignement et apprentissage. Sprimont (Mardaga).
- Fayol, M. & Jaffré, J.-P. (2008): Orthographier. Paris (PUF).
- Fayol, M. & Largy, P. (1995): Une approche cognitive fonctionnelle de l'orthographe grammaticale. In: Langue française, Volume 95, n°95, 80-98.
- Pacton, S., Fayol, M. & Perruchet, P. (1999): L'apprentissage de l'orthographe lexicale: le cas des régularités. In: Langue française, Volume 124, n°124, 23-39.
- Salins, G.-D. (1996): Grammaire pour l'enseignement/apprentissage du FLE. Paris (Didier-Hatier).

Pour une méthode active de remédiation orthographique pour des apprenants allophones

Claude GRUAZ

Teaching orthography to allophone adults is the source of specific problems to which the knowledge of grammatical rules applied to exercises gives no satisfactory answer. These persons have a competence of their own language and the question is how to let them acquire new automatisms.

The progression put forward is:

1. Identification of the deviations from the rules by means of a diagram which helps to draw an orthographic profile.
2. Orthographical profile.
3. Remedying, taking into account linguistic considerations and the competence of the person.
4. Pedagogical processes based on analogy and substitution.
5. Positive evaluation.

La démarche scolaire d'apprentissage de l'orthographe consiste à enseigner des règles, dans une progression qui va des plus simples aux plus compliquées, et de procéder ensuite à des exercices d'application. Elle ne peut s'appliquer à des apprenants étrangers qui ont une compétence orthographique dans leur propre langue et possèdent des automatismes. Chaque cas sera particulier dans la mesure où la langue de départ sera différente¹.

La démarche qui fait l'objet de cet article se fonde sur l'expérience acquise lors d'une recherche à la fois théorique et pratique menée avec des formateurs auprès de personnes ayant des difficultés en orthographe².

Il ne s'agit pas de procéder à une démarche de transfert qui consisterait d'abord à repérer ce que les règles de la langue de départ et le français ont en commun et en quoi elles se distinguent les unes des autres, puis à établir des correspondances entre ces règles.

La démarche générale est schématisée dans la figure 1.

Le principe général est de faire acquérir à l'apprenant une stratégie d'ajustement au contexte, voire de contournement, c'est-à-dire des réflexes qui lui permettront de construire une phrase dans laquelle ses éventuelles lacunes n'apparaîtront pas.

¹ Ce texte prend en compte les *Rectifications de l'orthographe* parues au *Journal officiel* (Documents administratifs) du 6 décembre 1990.

² Crefor (2009), Goffinet (2010).

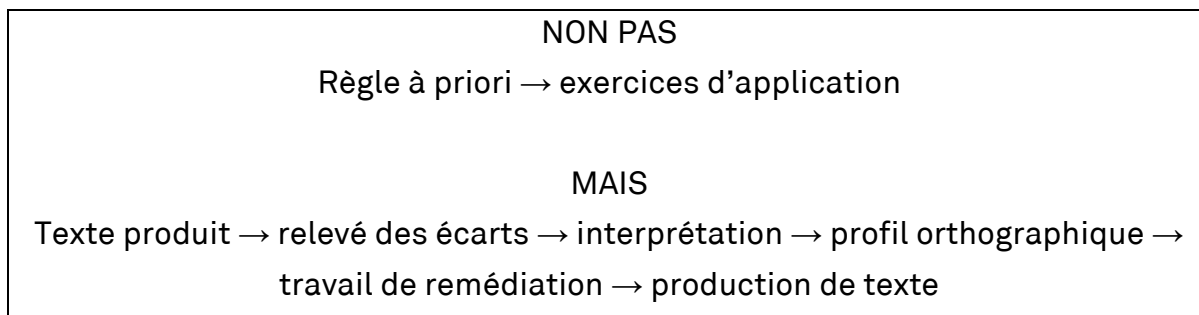


Fig. 1. Démarche générale.

Nous examinerons successivement:

- l'identification des écarts,
- le profil orthographique,
- la remédiation,
- l'analogie,
- l'évaluation positive.

1. Identification des écarts

La démarche a pour point de départ un texte écrit par l'apprenant.

Les écarts sont dus à une maîtrise insuffisante de l'oral ou proprement graphiques.

1.1 *Écarts dus à des erreurs de prononciation.*

Avant toute chose, il convient de connaître et d'identifier les difficultés des apprenants, c'est-à-dire de repérer les écarts qu'ils produisent dans des textes qu'ils écrivent en français³.

Une grande part des écarts est due à des erreurs de prononciation. Celle-ci constitue une difficulté pour les apprenants dont la langue d'origine peut différer par le nombre et par la nature des sons. Par exemple, l'arabe possède trois voyelles, lesquelles peuvent être longues ou brèves, alors que l'allongement n'est plus guère distinctif en français standard.

Les sons du français se répartissent en 11 voyelles orales, 3 semi-voyelles et 17 consonnes.

Une première difficulté est due au fait que certaines voyelles ont un double timbre, ex. *par*[a] / *pas*[ɑ], *beau*[o] / *bol*[ɔ], etc.

³ Nous éviterons autant que faire se peut d'employer des termes techniques, à l'exception de « graphème ». Pour une présentation plus théorique, on pourra se reporter à Catach (1980) et à Gruaz (1990).

Une seconde difficulté repose sur la distinction entre consonnes sourdes et consonnes sonores, ex. *pas* [p] / *bas* [b], *fais* [f] / *vais* [v], *ton* [t] / *don* [d], etc.

Lorsqu'un écart graphique porte sur une graphie liée à l'oral, il faut s'assurer prioritairement que le son correspondant est prononcé correctement ou, à tout le moins, que l'apprenant perçoit la différence entre les deux sons.

1.2 Écarts graphiques

Une grille à double entrée (Fig. 2 en annexe) permet d'identifier les écarts graphiques⁴.

En colonne sont indiqués les modes d'écart:

- confusion, ex. *ansuite* pour *ensuite*,
- omission, ex. *les déjeuner* pour *les déjeuners*,
- ajout, ex. *le travaille* pour *le travail*,
- déplacement, ex. *logn* pour *long*.

En ligne figurent les types d'écart:

- dessin de la lettre,
- écriture du son,
- catégorisation grammaticale,
- nom et adjectif,
- verbe,
- marques,
- le mot et la phrase.

(Nous renvoyons à la grille 2 ci-dessous, les exemples retenus dans ce qui suit se situent, par souci de simplification, uniquement dans la colonne "confusion"; certaines cases contiennent des exemples possibles mais non relevés dans des textes).

1.2.1 Le dessin de la lettre

Les lettres peuvent être mal formées (ex. *n* dans *lovg* pour *long*), en particulier lorsque les caractères de la langue L1 sont différents (cyrilliques, arabes, etc.). C'est parfois le cas des majuscules (ex. V pour U;

⁴ Cette grille repose sur le plurisystème graphique de Nina Catach (1980) et la grammaire homologique de C. Gruaz (1990). La terminologie a été volontairement simplifiée.

minuscules à la place des majuscules ou l'inverse), qui peuvent aussi être mal placées (ex. uN aVion).

1.2.2 L'écriture du son

L'unité de transcription du son n'est pas la lettre mais le graphème, unité fonctionnelle minimale de la chaîne écrite, ex. *b*, *eau*, *x* dans *beaux*.

Le français a 130 graphèmes⁵ qui transcrivent 36 phonèmes⁶, soit en moyenne 3,5 graphèmes pour un même son. Le risque d'erreur est donc grand dans le domaine de la transcription d'un son, auquel il convient d'ajouter les graphèmes qui ont une autre fonction, grammaticale ou distinctive (voir ci-dessous).

La relation de l'oral à l'écrit revêt trois aspects:

- a) la prononciation et la transcription sont incorrectes; la rectification ne peut se faire qu'après avoir entendu l'apprenant; la remédiation doit alors porter prioritairement sur l'oral;
- b) la prononciation est correcte mais la transcription ne l'est pas; ce qui est en cause ici est la perception ou la reconnaissance du son, ex. [p] de *par* confondu avec [b] de *bar*;
- c) la prononciation est correcte et la transcription est possible mais incorrecte, ex. *ansuite* pour *ensuite*; la remédiation fait appel aux régularités lexicales ou grammaticales, lorsqu'elles existent.

1.2.3 L'aspect morphologique

Une difficulté majeure est de reconnaître la catégorie des mots, surtout, mais non exclusivement, la distinction entre les nominaux et les verbes. Les nominaux peuvent être confondus avec des verbes, ex. *il *travail* pour *il travaille* ou l'inverse, ex. *le *déjeunai* pour *le déjeuner*.

Pour les noms et les adjectifs, la confusion porte sur les marques de genre (ex. *la *bel histoire*) ou de nombre (ex. *les *animaus*).

Pour les verbes, la confusion apparaît dans la conjugaison (ex. *nous étiont*, avec un *t* de troisième personne, pour *nous étions*). Elle peut aussi porter sur le nombre (ex. *ils *cours* pour *ils courent*) ou le genre (ex. *elle est *devenu* pour *elle est devenue*).

1.2.4 Les marques spécifiques

Des marques spécifiques rejoignent les domaines grammatical et lexical.

⁵ Cf. Catach (1980: 10-15).

⁶ Ibid, p. 37.

a) *Marques de rappel de féminin*

La présence de lettres muettes, généralement en finale de mot, peut être détectée par le recours au féminin, ex. *t* final de *petit* prononcé dans *petite*. C'est ainsi que l'on peut rencontrer *un moment *bénit* (au lieu de *béni*) à cause de *bénite*.

b) *Marques de rappel des dérivés*

Des lettres muettes, finales ou internes, sont également prononcées dans les dérivés, ex. *p* final de *champ* prononcé dans *champêtre*, *p* et *s* prononcés dans *corpuscule*, *t* final de *petit* prononcé dans *petitesse*; ce *t* est donc une marque de rappel à double titre, en tant que marque grammaticale et lexicale. La graphie **verd* serait déduite de *verdeur*.

c) *Marques distinctives*

Ces lettres servent à distinguer des mots qui ont la même prononciation, ex. le *c* de *sceau* différencie ce mot de *seau*. L'absence du *h* de *thon* entraîne la confusion avec *ton*.

d) *Marques de liaison*

Une lettre finale muette peut être prononcée en position de liaison avec le début du mot suivant à initiale vocalique, ex. *s* de *grands* est prononcé dans *les grands animaux* mais ne l'est pas dans *les grands fauves*. Lorsque cette lettre n'a pas de valeur grammaticale, elle est maintenue devant un mot à initiale consonantique, ex. *s* de *moins* prononcé dans *moins aride*, non prononcé dans *moins sec*; la graphie **moin sec* peut être attribuée à l'ignorance de cette régularité. Notons que la graphie de certaines lettres de liaison ne correspond pas à leur prononciation ex. *d* de *prend* est écrit *t* dans **prent-il*.

Remarque: une même lettre peut avoir plusieurs fonctions, ex. le *t* final de *toit* est une marque de rappel dérivative (*toiture*) et une marque distinctive (*toit / toi*).

La distinction est globale lorsque la graphie distingue des mots qui se prononcent de la même façon, ex. *seau*, *sceau*, *sot*.

1.2.5 Les lettres hors système

Certaines lettres n'ont plus de fonction et sont parfois oubliées, elles sont de nature étymologique, ex. *h* de *humeur* (latin *humor*) ou relèvent de l'histoire de l'orthographe, ex. *h* de *huit* (latin *octo*) ou *huile* (latin *oleum*).

1.2.6 Le mot

a) *Composants du mot*

Certaines parties du mot font l'objet d'erreurs, ex. le suffixe *eau* de *renardeau* écrit *au* comme le graphème final de *landau*.

b) *Segmentation du mot*

Lorsqu'un mot n'est pas identifié ou est ignoré, la segmentation de la séquence dans laquelle il figure peut être erronée, ex. *à la vance* pour *à l'avance*.

1.2.7 La phrase et la ponctuation

La ponctuation peut ne pas traduire la structure de la phrase, ex. *Jean et Marie; travaillent*.

2. Profil orthographique

En faisant la somme des écarts figurant dans chaque ligne, on repère les types qui sont les plus fréquents, par exemple tous les cas de prononciation correcte mais de transcription incorrecte.

De la même façon, la somme des écarts contenus dans les colonnes indique si les modes les plus récurrents sont la confusion, l'omission, l'ajout ou le déplacement.

Les cases situées à l'intersection des lignes et des colonnes informent précisément sur les écarts les plus souvent rencontrés, par exemple confusion des verbes avec les noms et les adjectifs.

Mais il s'agit là de données strictement quantitatives et l'on se gardera d'en déduire que l'on doit commencer la remédiation par celle des écarts les plus fréquents. Avant de procéder à la remédiation, il convient de procéder à une analyse linguistique et à des échanges avec l'apprenant afin d'affiner ces observations. Ces échanges se poursuivront tout au long du travail.

3. Remédiation

L'objectif du processus de remédiation est de donner à l'apprenant de véritables réflexes orthographiques qui lui permettront de surmonter les difficultés qu'il rencontre et qui se substitueront à la fonction que les règles ou les régularités, voire les exceptions, enseignées à l'école étaient supposées assurer car, pour autant qu'il les connaisse, il faudrait qu'il sache laquelle appliquer dans un contexte donné.

Dès le début, la démarche doit reposer sur le vécu de l'apprenant. Le travail de remédiation portera donc sur un texte dont il est l'auteur.

Une indication essentielle est la conscience que l'écrivain a de ses difficultés. Cette indication vient en complément des informations données par la grille, laquelle a pu faire apparaître des lacunes dont l'apprenant n'avait pas toujours conscience et dont l'observateur aurait pu ne pas mesurer la récurrence. Un entretien préalable fera apparaître ses propres représentations de l'orthographe en général ("c'est incompréhensible", "je n'y arriverai jamais") et l'existence de blocages personnels en face de certaines difficultés. Et tout écrivain francophone n'est-il pas en situation d'éternel apprenant devant un accord du participe passé, devant le doublement de consonnes ou devant certains emplois du *x* de pluriel?⁷

Les données fournies par le profil orthographique et les représentations de l'écrivain sont alors mises en relation avec le système orthographique, ce qui permettra au formateur d'établir un ordre de priorité entre les écarts. Ainsi, l'approche des expressions graphiques du nombre ne peut être traitée de façon satisfaisante si la distinction entre les nominaux et le verbe n'est pas maîtrisée.

Précisons que par système, s'entend non pas un ensemble de règles strictes, toujours vraies et établies une fois pour toutes, mais un ensemble de régularités, généralement vraies, mais qui, dans l'état actuel de l'orthographe, s'accompagnent de trop nombreuses exceptions. Ce sont donc ces régularités qu'il conviendra de faire acquérir en priorité. Ainsi écrire *les *chevaus* est bien davantage une "faute" vis-à-vis de la norme que du système tel qu'il devrait être et de la logique.

Lors de la remédiation, l'apprenant doit être actif: il s'agit essentiellement de lui faire acquérir des moyens de résoudre ultérieurement par lui-même les problèmes qu'il rencontre lorsqu'il est seul devant la page.

La remédiation prend la forme d'un dialogue suivi entre l'apprenant et le formateur. Ce dialogue est fondé à la fois sur le contenu strictement linguistique tel que le formateur l'a envisagé et sur les représentations de l'apprenant. Le formateur doit être prêt à modifier le processus engagé en fonction des réactions de l'apprenant.

Sur quoi repose ce dialogue?

Le principe fondamental est de partir des compétences de l'apprenant. Quel que soit son niveau orthographique en français, il a une compétence, à tout le moins dans sa langue maternelle. Il ne s'agit donc pas de lui apporter des connaissances entièrement nouvelles "de l'extérieur", mais de détecter ses connaissances réelles, "intérieures" en quelque sorte, et de

⁷ Cf. Gruaz (dir.): (2009, 2010).

développer celles-ci. Par exemple, un apprenant anglais saura ce qu'est un accord et un adjectif, mais il pourra éprouver des difficultés à accorder un adjectif en français.

C'est sur le repérage des compétences réelles, qui doit précéder toute remédiation, que se construit une hiérarchisation des sujets traités. La remédiation linéaire, progressant du début à la fin du texte, ne saurait en effet être retenue.

On peut penser que l'apprenant qui écrit *les hibous dormes* connaît:

1. la notion de pluriel,
2. le principe de l'accord entre le sujet et le verbe,
3. le *s* comme marque de pluriel.

mais qu'il ne maîtrise pas :

1. les marques de pluriel des verbes,
2. que le pluriel de *hibou* s'écrit avec un *x* et non un *s*... et il reste au formateur d'espérer que l'apprenant ne lui demande pas pourquoi⁸.

Une fois repérée, une lacune doit être replacée dans le système – au sens que nous donnons à ce terme – orthographique français: chaque point est à l'intersection de plusieurs voies à l'intérieur d'un réseau de régularités. Selon les compétences de l'apprenant, une voie sera privilégiée. Ainsi la remédiation portant sur la graphie *déjeunai* mise pour *déjeuner* peut prendre appui sur:

1. le contexte syntaxique: l'examen de contextes produits pour cette étude permet de découvrir que *déjeunai* sera le plus souvent précédé d'un pronom personnel et variera selon les personnes, les temps et les modes verbaux, alors que *déjeuner* sera précédé d'un article (*le déjeuner*) ou d'une préposition (*Je viens de déjeuner*). Selon la compétence métalinguistique de l'apprenant, le recours à une terminologie grammaticale sera plus ou moins efficace. Toutefois, une démarche très marquée par son caractère métalinguistique ne semble pas devoir être privilégiée.
2. le rapport à l'oral: la substitution du verbe *déjeuner* par un verbe du 2^e ou du 3^e groupe dans le même contexte écartera l'homophonie (*Je viens de déjeuner / Je viens de courir*).

La référence au système doit être préférée au recours au sens: on sait que la notion de pluriel ne doit pas être confondue avec celle de pluralité; de ce fait il est plus pertinent de justifier l'accord d'un verbe par le déterminant qui précède le nom sujet que par son sens propre, par exemple *ténèbres* est un pluriel grammatical et un singulier sémantique (*ténèbres* a le sens

⁸ Il faudrait alors lui préciser que l'Académie a renoncé à appliquer ce qu'elle avait accepté en 1908!

de *nuît*) précédé d'un déterminant au pluriel dans *les ténèbres tombent* et par voie de conséquence le verbe est au pluriel.

Le travail de remédiation tel qu'il est présenté ici est individuel, mais il pourrait aussi faire l'objet d'un travail de groupe, le formateur devant alors préalablement procéder à une synthèse des écarts observés dans les profils des apprenants et entreprendre une réflexion collective. Ce travail collectif peut présenter l'avantage de faire disparaître les blocages qui existent chez certains apprenants.

4. L'analogie

La démarche de remédiation préconisée repose essentiellement sur le principe d'analogie.

Ce principe est le suivant: le remplacement d'un élément (mot ou lettre) – le déclencheur –, par un autre dans une phrase a pour effet de faire apparaître

- soit la graphie correcte de l'élément remplacé : dans la phrase *Je vais arriver demain*, va-t-on écrire *arriver* ou *arrivé*? il suffit de remplacer *arriver* par *venir*.
- soit la graphie correcte d'autres éléments présents dans la séquence considérée. Il suffit alors de remplacer le terme qui pose problème par un autre : va-t-on écrire *un visage commun* ou *commin*? on remplace *visage* par *figure*, soit *une figure commune*, le *u* prononcé dans ce dernier mot justifie la graphie *un*. Pour reprendre l'exemple précédent, on peut remplacer *je vais* par *j'ai* ou *je suis*, et l'on écrira alors *arrivé* dans *je suis arrivé* par analogie avec *venu* dans *je suis venu* et *arriver* dans *je vais arriver* par analogie avec *venir* dans *je vais venir*⁹.

La remédiation orthographique est donc essentiellement une manipulation de la langue, non seulement écrite mais aussi orale. Le rôle du formateur est d'entraîner l'apprenant à procéder à ces substitutions, ce qui aura pour effet indirect de réactiver des connaissances dont il n'a peut-être pas conscience, des "connaissances inconnues" en quelque sorte. L'acquisition de tels automatismes est au cœur du processus de remédiation.

5. L'évaluation positive

Tout au long du processus de remédiation, le formateur soulignera ce qu'il y a de positif dans la production écrite de l'apprenant. Il est essentiel de lui montrer qu'il connaît de l'orthographe sinon l'orthographe, que

⁹ Crefor (2009: 37ss.).

l'orthographe n'est pas un domaine auquel il n'a pas accès mais dans lequel il a d'ores et déjà pénétré.

Le problème pour lui est d'en découvrir les grandes avenues, les grandes régularités et non pas les multiples ruelles et exceptions qui l'encombrent¹⁰ et sur lesquelles l'école n'a que trop tendance à mettre l'accent (à insister sur les sept pluriels en *oux*, on en vient à mettre un *x* à des mots nouveaux, comme en témoigne le titre du film *Les ripoux*).

Conclusion

La participation active de l'apprenant est essentielle à l'acquisition de l'orthographe. Le rôle du formateur n'est donc pas d'enseigner des règles, mais d'habituer l'apprenant à résoudre par lui-même les problèmes qui se poseront à lui lorsqu'il ne sera plus accompagné.

Pour atteindre cet objectif, le formateur doit disposer de moyens qui lui permettront dans un premier temps d'identifier les difficultés réelles de l'apprenant et dans un second temps de les surmonter.

Le recours à l'analogie sera bien souvent d'une grande efficacité dans la mesure où cette analogie fera appel aux connaissances de l'apprenant, qu'il s'agisse de sa propre langue ou de la langue maternelle pour un apprenant allophone.

Bibliographie

- Catach, N., avec la coll. de Gruaz, C. & Duprez, A. (1980): L'orthographe française, traité théorique et pratique. Paris (Nathan).
- Crefor Haute-Normandie, coordination Mercier J.-P., direction Gruaz, C., collaboration Bottois, F., Chesnel, C. & Funkiewiez, F (2009): Recherche-action Orthographe et illettrisme.
(http://www.crefor-hn.fr/sites/default/files/Recherche_action_orthographe.pdf)
- Goffinet, S.- A. (2010): Orthographe et illettrisme. In: Le journal de l'alpha, N° 176, novembre 2010. Bruxelles (Lire et écrire Communauté française).
- Gruaz C. (1990): Du signe au sens, pour une grammaire homologique des composants du mot. Rouen (Presses Universitaires de l'Université de Rouen).
- Gruaz C. (dir.) (2009): Les consonnes doubles, féminins et dérivés. Limoges (Lambert-Lucas).
- Gruaz C. (dir.) (2010): Le x final. Limoges (Lambert-Lucas).

¹⁰ On pourra sur ce sujet consulter le site de l'association ÉROFA, *Études pour une Rationalisation de l'Orthographe Française d'Aujourd'hui*, <http://erofa.free.fr>.

Annexe 1: Grille d'évaluation diagnostique

	Confusion	Omission	Ajout	Déplacement
Dessin de la lettre				
Erreur	lovg <i>pr</i> long	long <i>pr</i> lomg	lomg <i>pr</i> long	
Majuscule	Uous <i>pr</i> Vous	nicole <i>pr</i> Nicole	je Vais <i>pr</i> je vais	je vAis <i>pr</i> je vais
Ecriture du son				
Prononciation et transcription incorrectes	(il faut connaître la prononciation)	(il faut connaître la prononciation)	(il faut connaître la prononciation)	(il faut connaître la prononciation)
Prononciation correcte mais transcription incorrecte	dérole <i>pr</i> déroule croupe <i>pr</i> groupe	cuit <i>pr</i> cuite	achetrer <i>pr</i> acheter	fatie <i>pr</i> faite
prononciation correcte et transcription possible mais incorrecte	ansuite <i>pr</i> ensuite	lomg <i>pr</i> long	longt <i>pr</i> long	logn <i>pr</i> long
Catégorisation grammaticale				
Nom/adjectif pour verbe	il travail <i>pr</i> il travaille	tu travail <i>pr</i> tu travailles	tu travaux <i>pr</i> tu travailles	
Verbe pour nom/adjectif	le déjeuner <i>pr</i> le déjeuner	les déjeuner <i>pr</i> les déjeuners	le travaille <i>pr</i> le travail	
Nom et adjectif				
Genre des noms et adjectifs	la bel histoire <i>pr</i> la belle histoire	venu <i>pr</i> venue	venue <i>pr</i> venu	
Nombre des noms et adjectifs	les animaus <i>pr</i> les animaux	les achat <i>pr</i> les achats	nouveaux <i>pr</i> nouveau	
Verbe				
Nombre	ils cours <i>pr</i> ils courent	il cour <i>pr</i> il court	il manget <i>pr</i> il mange	
Conjugaison	nous étiont <i>pr</i> nous étions	il es <i>pr</i> il est	tu est <i>pr</i> tu es	
Genre	elle est devenu <i>pr</i> devenue	elle est devenu <i>pr</i> devenue	elle est devenue <i>pr</i> devenu	
Marques				
De féminin	un moment bénit/bénite <i>pr</i> bénite	gran <i>pr</i> grand (grande)	une grand fille <i>pr</i> grande	
De dérivés	verd <i>pr</i> vert (verdure)	gran <i>pr</i> grand (grandeur)	abrit <i>pr</i> abri (abriter)	
Distinctives	ton <i>pr</i> thon	travail <i>pr</i> travaille	travaille <i>pr</i> travail	
De liaison	moin sec	le animaux <i>pr</i> les animaux	Des zanimaux	le sanimaux <i>pr</i> les animaux
Lettres hors système		umeur uile		
Le mot et la phrase				
Composants du mot	renardau <i>pr</i> renardeau	comme <i>pr</i> commence	comment <i>pr</i> comme	fréquememnt <i>pr</i> fréquemment
Segmentation du mot et apostrophe	ala vance <i>pr</i> à l'avance	alavance <i>pr</i> à l'avance	par'avent <i>pr</i> paravent	
Signes de ponctuation	Jean et Marie; travaillent	Jean Marie et Pierre travaillent	Jean et, Marie travaillent	Jean et Marie, travaillent

Pour une analyse morphologique des productions écrites d'élèves FLS

Marie-Claude LE BOT & Elisabeth RICHARD

EA LIDILE 3874, UFR Langues, Université Rennes 2

In this article we study a corpus of written work by allophonic secondary school pupils (aged 11-15), the majority of which are Turkish native speakers whose language at school is French.

Although their written work in French does not conform with conventional spelling due to the fact that it is based on sound, the linguistic description of the corpus also shows that it does not comply either with phonetic spelling. Their written work both testifies to a number of skills acquired through learning and provides an explanation as to the origin of their spelling mistakes. It is our contention that the strategies of French language spelling are for the most part based on the paradigmatic dimension of the language, independently of its segmental dimension structuring its verbal or nominal units.

Notre réflexion s'appuie sur l'examen d'un corpus d'élèves d'un collège d'une grande agglomération bretonne¹. Le corpus a pour particularité de retenir les productions écrites d'élèves allophones, majoritairement turcophones, qui ont pour langue de scolarité le français. Les performances langagières en français de ces élèves présentent une disparité très nette entre production orale et production écrite. En effet, si la performance orale de ces élèves est satisfaisante, en revanche leur performance écrite est un problème majeur et a pour effet de les mettre en grande difficulté scolaire.

En début d'année scolaire, les responsables du collège ont repéré les 24 élèves allophones les plus en difficulté et ont souhaité vérifier leur niveau de langue en les soumettant à une évaluation de niveau A2². Homogène du point de vue des productions langagières, le groupe ne l'était pas du point de vue scolaire puisqu'il était composé d'élèves régulièrement inscrits dans des classes de 6^{ème}, 5^{ème} et 4^{ème} du collège (âgés de 11 à 15 ans). Plus précisément, le groupe était organisé comme suit: 14 élèves inscrits en

¹ Le corpus a été recueilli dans le cadre du Master "Etudes linguistiques appliquées à l'enseignement et à l'apprentissage des langues", Université Rennes 2. Que soient ici chaleureusement remerciés les différents acteurs du collège d'accueil et les étudiants stagiaires du Master.

² Les épreuves se sont déroulées le 10 décembre 2010.

6^{ème}, dont 6 en section SEGPA³ (11 garçons, 3 filles); 5 élèves en 5^{ème} (4 garçons, 1 fille); 4 élèves en 4^{ème} (1 garçon, 3 filles).

Cette étude prend appui sur un corpus composé des 48 productions écrites dirigées obtenues dans le cadre de cette évaluation et correspondant à deux exercices différents⁴. Pour le premier exercice, les élèves étaient invités à écrire le récit d'un voyage. La consigne écrite était accompagnée de quatre photographies dont trois étaient sous-titrées. Le second exercice – de type épistolaire – consistait à répondre à une lettre d'invitation accompagnant la consigne. Il s'agissait, dans ce cas, de décliner l'invitation et de s'en expliquer.

Pour chacun des exercices, les élèves devaient produire un texte de 60 à 80 mots et cet aspect de la consigne a été bien respecté.

L'objectif de l'étude que nous présentons est d'engager une réflexion qui puisse, à long terme, répondre aux interrogations de l'équipe enseignante du collège devant des erreurs récurrentes, mais 'inclassables', de ces élèves allophones lorsqu'on les rapporte aux performances écrites des autres élèves de ces classes. L'examen du corpus dont nous présentons ici l'analyse linguistique est une première étape qui devrait ouvrir un certain nombre de pistes de réflexion méthodologiques et didactiques.

Nous avons abordé ces productions en appliquant la méthodologie de l'analyse d'erreurs, dont l'objectif n'est pas de simplement mesurer des écarts par rapport à une norme, mais d'évaluer aussi les acquis et les stratégies d'apprentissage qui sont à l'œuvre. En cela, nous confortons l'hypothèse développée notamment par D. Cogis (2002: 37): "Quoi qu'il en soit, la disparité supposée des productions graphiques des enfants – "le n'importe quoi" – se réduit singulièrement, si l'on interprète leurs graphies en tant que produits résultant de certaines procédures". Notre démarche s'accorde avec celle utilisée pour la description des interlangues d'apprenants, qui rapporte ces productions particulières à des actualisations de systèmes intermédiaires⁵, spécifiques et caractéristiques d'un apprentissage en train de se faire.

Pour autant, et c'est ce qui en fait leur spécificité, les difficultés linguistiques que présentent ces élèves à l'écrit ne sont pas assimilables à

³ SEGPA: Section d'Enseignement Général et Professionnel Adapté destinée aux élèves "présentant des difficultés d'apprentissage graves et durables".

<http://eduscol.education.fr/pid23266-cid46765/sections-d-enseignement-general-et-professionnel-adapte.html>

⁴ Il s'agit de productions écrites à la main et non sur traitement de texte. L'usage du dictionnaire n'était pas autorisé.

⁵ "La double caractéristique de l'interlangue repose d'une part sur [...] la nature et les règles de grammaire intériorisée, d'autre part sur son caractère évolutif". Besse, H. & Porquier, R. (1984: 217).

celles des publics allophones de FLE et l'analyse des productions amène à déconstruire l'évaluation globalement déficitaire qu'elles suscitent au premier abord. Une analyse des systèmes graphiques mobilisés permet d'affiner l'évaluation des acquis en matière d'écriture et de tracer en creux ce qui n'a pas fait l'objet d'un apprentissage pertinent.

Notre exposé rendra compte de la méthodologie que nous avons choisie. Nous nous attacherons à relever les marques d'une acquisition effective des apprentissages scolaires, en ce qui concerne la construction du texte, d'une part, et l'exploitation des systèmes graphiques d'autre part. Cette description nous permettra ensuite de cerner au plus près, ce qui, selon nous, constitue le déficit majeur des productions écrites de ces élèves et qui concerne la morphologie nominale et verbale.

1. La construction des textes

Bien que les textes que nous reproduisons ci-après ne correspondent pas, loin s'en faut, aux productions attendues d'élèves de collège, elles n'en montrent pas moins une maîtrise certaine – mais partielle – de la langue française et de son écriture.

Les deux textes qui suivent sont des productions complètes et nous avons maintenu la mise en page adoptée par les élèves eux-mêmes. Le premier est un récit de voyage, le second, une lettre. Nous examinerons les écrits du point de vue textuel afin de faire le point sur les acquis scolaires en matière de genre, de cohésion et d'argumentation.

(1) Récit de voyage

A Paris je suis monter a la touR efféle, je suis paRti a la fête foRaine.
Je voyeR tout de la touR efféle. Aussi ans 2000 il y a vais un feudartifise est ya un
Pétare qui ya explosé a coté de moi j'ai eu
tellemen peur que j'ai sur sote, j'aimé
bien voir les avion voleR en laiRs il on
avais beaucoup d'avions. Paris s'est beau
mes il ya beaucoup de brux s'est les avions
si tu ten doR les avions te raivaille, on
n'auras plus besoin de Réveille.
Fin
Kemal

(2) Lettre

Bonjour, merci pour votre invitation Philippe mait
Malheureusement je ne peut pas venir, quar je d'évrai
partir en tunisie mait je pourrait petêtre venir lanné
prochaine avec toi et ce que ta famille va bien
ta sœur ton frère et tout ta famille est passe leur
bonjour aurevoir.
merci pour l'invitation
Chaouki

1.1 *Les genres textuels*

Dans un premier abord, c'est bien l'aspect très désorganisé de ces textes qui est le plus remarquable, et cette appréciation tient beaucoup au caractère relativement peu prédictible des séquences graphiques qui se juxtaposent sans prise en compte conventionnelle des marques graphiques de la ponctuation et des majuscules. Mais, alors que ces textes se présentent, dans un premier temps, comme proprement 'illisibles', il s'avère que ces mêmes écrits subissent avec succès le test de lecture à voix haute (serait-il plus exact de parler de test de 'déchiffrement'?). En effet, dans leur version oralisée, la majorité d'entre eux n'engendre que très peu de problèmes d'interprétation dans la mesure où, non seulement ils répondent à la consigne imposée, mais ils satisfont aussi aux contraintes principales de l'organisation générale des différents genres de textes. C'est ainsi que pour le premier écrit, les réponses sont conformes à ce qui était attendu, c'est-à-dire un récit qui intègre des éléments visuels imposés, ainsi que leurs annotations toponymiques. De la même manière, le genre épistolaire attendu est convoqué de manière adéquate pour répondre au second exercice, et les textes intègrent les conventions linguistiques du genre (termes d'adresse et de politesse, salutations, mise en page).

1.2 *La cohésion*

Pour ce qui est de l'analyse textuelle des productions, on note que les élèves exploitent de manière pertinente des marqueurs formels qui assurent la cohésion des textes. En effet, la majorité des écrits atteste d'une structuration interne et d'une organisation chronologique des événements qui impliquent une maîtrise des oppositions temporelles (passé-composé *vs* imparfait pour le récit, présent *vs* futur pour la lettre), des marqueurs lexicaux adéquats ainsi qu'une exploitation conforme du système anaphorique pronominal ou lexical, comme le montrent les exemples ci-dessous:

(3) est jesuis rentré chez moi pour dormir quart j'étais fatiger parceQue tout ce que j'ai fait c'était beaucoup

(4) Après avoir finis la visite. on est aller voir le film d'harry potter

(5) Le dix-huit juiellet j'étaient à Paris. le premier jour j'ai visité le château de versaille et ensuite avec mes parents nous sommes aller manger des glaces dans un parc. L'avant d'ernier jour nous avons visité la cathédrale de paris [...]

(6) mais je pourrait petêtre venir lanné prochaine avec toi

(7) je suis conten que tu me inviter chez toi, malheureusement je ne peu pas venir à pariscar je pars en Amérique pendent 2 mois excuse moi.

1.3 *L'argumentation*

Nous ferons la même remarque en ce qui concerne la maîtrise du système argumentatif, puisque les productions attestent de capacités de hiérarchisation des arguments (système oppositif et explicatif), de mise en forme de l'information (formes disloquées, présentatifs), d'expression de la subjectivité (intensité, marquage énonciatif), comme l'illustrent les extraits suivants:

(8) cette année j'ai pas beaucoup des vacances, par còntre l'année prochaine j'aurais tous les vacances d'été

(9) mait Malheureusement je ne peut pas venir

(10) j'ai aussi visitér le château de versaille

(11) Après se que j'ai particulièrement aimer ses le futuroscope

C'est ainsi que l'analyse textuelle conduit à nuancer l'impression immédiate donnée par une première lecture des productions écrites de ces élèves "en très grande difficulté scolaire". En effet, ces écrits témoignent d'une exploitation pertinente des ressources de la langue française en matière d'expression de la temporalité, de l'argumentation et de la subjectivité, conformes, en grande partie, aux attentes scolaires correspondant à leur niveau d'étude.

Cette observation, permet donc de pondérer d'une manière positive l'évaluation de ces élèves, qui ont acquis un certain nombre d'outils linguistiques et qui sont effectivement en phase d'acquisition des compétences argumentatives et lexicales dispensées au collège.

Cette première conclusion est par ailleurs cohérente avec la remarque déjà faite que ces élèves ne présentent pas de productions orales déficitaires majeures, et il est clair que l'appui sur une langue orale performante joue ici un rôle non négligeable.

Pour autant, l'échec scolaire est bien le dénominateur commun de ce groupe d'élèves et nous montrerons que cette évaluation s'objective, en grande partie, sur l'aspect orthographique des productions écrites.

2. Les stratégies graphiques

Dans cette partie, nous nous intéresserons aux procédés de mise en lettres et nous porterons la description sur deux aspects des performances des élèves en distinguant les erreurs qui procèdent directement de leur condition d'élèves allophones de celles qui relèvent de *l'orthographe* au sens de Catach⁶.

⁶ Catach oppose la *graphie d'une langue* qu'elle définit comme "la manière d'écrire les sons ou les mots sans référence à une norme ou au système de langue", à *l'orthographe* qui est "la manière d'écrire les sons et les mots d'une langue en conformité avec le système de

2.1 *Les traces d'une langue première*

Un premier ensemble d'erreurs est à imputer à des phénomènes d'interférence avec la langue maternelle des élèves. C'est sans doute sur ce point que les performances ici décrites sont les plus proches de celles que l'on recueille auprès de publics FLE.

Plus précisément – dans un nombre relativement restreint de productions, toutefois – on relève des graphies erronées directement liées à un parasitage phonétique et phonologique, notamment sur la nasalité vocalique et le trait de voisement de certaines consonnes. Nous en donnons toutes les occurrences:

[ɑ̃] se trouve régulièrement graphié par "on", ou à l'inverse [ɔ̃] par "an, en"

(12) on a rontré chez ma tente

(12') Apres le matin en c'est reveiller en aprix le petit déjeuner.

[ɛ̃] est graphié par "an, en"

(13) an jardan; mersi de me an vite

(13') j'ai en petit peux eus la mal de mer

[y] graphié "i"

(14) visiter les batiments historique et des moniment

[d] graphié "t" ou, à l'inverse, [t] graphié "d"

(15) je vais de réserver un billet

(15') je toi aller en Allemagne

Ces erreurs sont régulières dans un apprentissage FLE dans la mesure où la graphie s'appuie sur les réalisations phonétiques orales des apprenants. Nous considérons qu'elles signalent, chez nos élèves, une maîtrise insuffisamment consolidée du système phonologique du français et permettent de poser l'hypothèse d'une exposition encore forte à la langue maternelle (notamment dans le cadre familial). Chez certains élèves, cette caractéristique FLE des productions apparaît aussi dans leurs difficultés à utiliser correctement les déterminants nominaux, qu'il s'agisse du choix du genre, ou de celui des allomorphes: "je suis allée avec mon famille", "on a regardai des plusieurs films". Il faut toutefois signaler que, dans notre

corpus, ces erreurs de type FLE occupent une place mineure et ne concernent que 6 élèves sur 24⁷.

2.2 *Les signes d'une stratégie ortho-graphique*

L'examen du corpus met au jour une seconde catégorie d'erreurs, massivement présente et généralisable à l'ensemble des écrits, cette fois. Sur ce point, notre corpus présente des différences très notables avec les tableaux d'erreurs relevés dans des corpus d'apprenants FLE, et contribue à faire de ces élèves allophones des représentants du public Français Langue Seconde (FLS),⁸ qui désigne, ici, ces élèves allophones qui évoluent dans deux communautés linguistiques relativement étanches, turcophone pour ce qui relève du contexte familial et francophone pour le reste de leurs échanges.

En effet, bien qu'erronées, les productions graphiques, ne sont pas pour autant aléatoires et les textes montrent une utilisation effective, mais erratique, de l'ensemble des procédés orthographiques qu'impose l'écriture du français. Même si les résultats ne sont pas ceux attendus, le projet graphique des élèves d'écrire du français en français est bien réel. En effet, les erreurs relevées attestent d'une connaissance très complète des diverses actualisations graphiques du français et signent la mise en œuvre d'une stratégie orthographique qui s'appuie sur le panel des règles en usage de la mise en lettres du français.

Autrement dit, pour ces élèves, le projet d'écriture de la langue s'alimente bien de la connaissance qu'ils ont d'un certain nombre de formes graphiques attestées en français. Cette remarque se voit vérifiée par deux caractéristiques de ces écrits: une correspondance phonie-graphie qui n'est pas aléatoire et des manifestations hypertrophiques de procédés graphiques spécifiques au français. Ces deux points viennent illustrer de manière complémentaire que ces élèves ont bien été engagés dans un processus d'acquisition de l'écrit et que leurs productions en portent les traces.

Il faut tout d'abord rappeler le caractère spectaculaire, parce que contrasté, de ces écrits qui font cohabiter – au sein d'un même texte et parfois dans même segment – la présence de séquences parfaitement orthographiées et d'autres qui ne le sont pas. Il est tout autant paradoxal de relever que les mots d'usage peu fréquent ne sont pas l'objet de plus

⁷ Dans le même ordre d'idées, on a noté qu'un élève ne maîtrise pas correctement la conjugaison du passé-composé, notamment pour ce qui est du choix de l'auxiliaire.

⁸ Cette étude confirme donc les projections de Cuq sur l'évolution de la population scolaire française qui prévoyait que " [la question du FLS], quasi réservée aux régions francophones hors la France métropolitaine, va être de plus en plus intégrée au débat scolaire français" (2005: 96).

d'erreurs que ceux qui ne le sont pas. Il semblerait même que ce soit l'inverse. Par exemple en (7), "malheureusement" et "excuse moi" sont correctement orthographiés à côté de "conten", "je ne peu pas" et "pendent", qui sont pourtant d'un usage courant.

Par ailleurs, l'analyse des productions erronées atteste que les procédés de mise en lettres des phonèmes sont appliqués et qu'ils aboutissent, majoritairement, à assurer une correspondance phonie-graphie. En effet, les séquences graphiées par les élèves correspondent à des séquences, potentiellement allographes, du français sans que ces graphies insolites n'altèrent pas la réalité phonique de ce qui est transcrit, comme en atteste le florilège suivant: "biyet", "biye", "ennuiller", "brier", "trenquile", "quar", "couzin".

De la même manière, l'exploitation du système des accents grave, aigu et circonflexe est à l'œuvre, souvent de façon non conforme, mais néanmoins efficiente ("le tramouè", "la fête forène"). On note cependant une manière radicale, et très fréquemment adoptée, de régler la question épineuse des accents: de très nombreux exemples montrent une exploitation systématique de la lettre "e" en lieu et place de la transcription de [e] ou [ɛ].

Pour autant, même si le qualificatif de 'phonétique' est tentant à utiliser pour spécifier ce mode d'écriture, il se révèle inexact parce qu'il suggère que la graphie de ces élèves procéderait d'un assujettissement exclusif à l'oral.

Or, il nous faut aussi prendre en compte un autre aspect de ces productions erronées qui touche les doubles consonnes, les lettres sans correspondant(s) phonique(s), les apostrophes, les accents. Il est à noter que toutes ces particularités propres à l'orthographe du français ont en commun d'échapper à une stricte stratégie de mise en correspondance phonie-graphie⁹. Leur actualisation dans la graphie du français mobilise d'autres connaissances, en particulier une analyse des niveaux morphologique, syntaxique et lexical de la langue. Une grande partie de l'effort demandé aux scripteurs du français lors de l'apprentissage porte sur ces procédés, et leur maîtrise en est, de ce fait, valorisée et très valorisante.

Nous relevons que les élèves n'ignorent pas tous ces aspects de la graphie du français, mais qu'ils en font un usage très inattendu, que nous pourrions qualifier de 'débridé', 'd'excessif'. Leurs choix graphiques ("bonjour philippe") renvoient l'image d'une écriture complexifiée, preuve qu'ils ont

⁹ De même, les observations de Fayol et Jaffré sur des élèves FLM à propos de "abhit, ihver": "Ces résultats suggèrent que les enfants n'acquièrent pas que les formes orthographiques en relation directe avec la phonologie. Ils apprennent aussi des indices orthographiques qu'ils placent parfois en position erronée, mais qui témoignent d'un apprentissage orthographique relativement indépendant de la phonologie." (2008:187)

bien intégré l'idée que l'écriture du français contraint de se plier à une *ortho*-graphie.

Il en est ainsi de l'exploitation exagérée des accents ou des doubles consonnes graphiques, comme le montrent ces exemples "par cône", "j'ai aussi visité", "chère philippe" et aussi "la semaine", "la deuxième semaine").

L'emploi abusif de l'apostrophe est également intéressant à observer car ce signe est très présent dans les écrits des élèves: "l'avant d'ernier jour", "mon cousin habitait j'uste à côté", "jusqu'que à 19h", "merci de t'on invitation". Si le résultat aboutit à une graphie inédite des mots concernés, on notera cependant que l'apostrophe vient se placer, dans la majorité des cas, derrière une lettre, ou une suite de lettres qui ne lui est pas inhabituelle en français: d'; j'; jusqu', t'.

Un autre aspect intéressant, et parmi les plus réguliers, est une exploitation – *a priori* erratique – de lettres sans correspondant phonique. Les erreurs alors rencontrées sont de deux types: soit il y a un mauvais choix de la lettre en question ("beaucoup de brux", "dans un restaurant", "une fois", les musique de Mozare), soit il y a un ajout sur un mot qui ne le demandait pas ("Mc Deaux", "un écran plasma", "un hôtele"). On note aussi que ces lettres ne résultent cependant pas d'un choix aléatoire, car elles sont effectivement exploitées comme morphogrammes (grammaticaux et/ou dérivatifs) du français (e, t, d, s, x) et nous n'avons pas rencontré de lettre 'insolite'.

Tout ceci confirme que ces élèves ont, non seulement, eu accès à un apprentissage de l'écrit du français, mais qu'ils en connaissent les formes valorisées sans être en mesure de les exploiter à bon escient. A ce titre, nous interpréterons ces erreurs comme des manifestations d'hypercorrection¹⁰, identifiée par P. Bourdieu¹¹, comme une marque de distinction sociale. Sur ce point, nous posons que le maniement correct des lettres sans correspondant phonique est le pendant de celui des liaisons facultatives dans la langue orale: là où leur actualisation à-propos dans les discours des locuteurs et/ou des scripteurs atteste d'une maîtrise de la norme officielle, une exploitation hypertrophiée est à la fois le signe d'une connaissance partielle de leur application et, dans le même temps, une reconnaissance de leur valeur. Cette remarque complète l'analyse que nous faisons de ces productions comme relevant d'une graphie qui inclut le

¹⁰ La définition qu'en donne Francard précise que "le concept d'hypercorrection [rend compte] de la propension de certains locuteurs à produire des formes qu'ils veulent conformes à un usage légitimé, mais qui, en réalité, s'en écartent", in Moreau (1997: 158).

Pour une discussion de la notion, cf. aussi Gadet (1997: 15).

¹¹ Bourdieu (1982: 84-85).

principe d'une orthographe, sans en maîtriser les règles, sinon de manière très parcellaire.

3. Un écrit désordonné

Il reste que ces écrits sont, en contexte scolaire, évalués très négativement et l'impression de très grande désorganisation qu'ils donnent à voir (à lire) n'est pas tempérée par les acquis dans l'apprentissage de l'écrit du français dont ils témoignent pourtant et que nous venons de présenter. C'est donc à une évaluation du déficit d'apprentissage que nous allons désormais nous intéresser, et les extraits suivants en sont de bons exemples:

(16) ensuite j'ai jouer avec mais jouer

(17) ma meRe ne ve pas car on a Reserver déplace pouR le Japon

(18) se te tres très Bien

(19) c'est ai trô beau, je tes bien la bas

(20) ont a regardait la télé

Ces séquences sont typiques du corpus et combinent deux caractéristiques qui altèrent gravement le processus de lecture de ces textes. Premièrement, elles s'insèrent dans un continuum graphique rarement affecté des marques de la ponctuation: pas de virgules, très peu de points, un emploi très fluctuant des majuscules, qui peuvent aussi bien apparaître au milieu d'une phrase comme au milieu d'un mot.

Deuxièmement, ces séquences graphiques alternent des suites de lettres et des blancs sans que les unes et les autres ne soient obligatoirement en correspondance avec ce qui est supposé être graphié, c'est-à-dire:

- un nom: (16) "avec mais jouer"; (17) "déplace"

- ou un verbe: (18) "se te"; (19) "c'est ai"; "je tes bien"; (20) "ont a regardait". Il est à noter que la graphie de l'unité verbale est plus massivement désorganisée que celle de l'unité nominale¹².

3. 1 *Un stock de formes écrites*

Nous postulons que la stratégie mise en œuvre s'appuie sur une connaissance de la langue ordonnée sur le mode d'un stock de mots et dont l'écriture a fait l'objet d'un apprentissage et d'une mémorisation pour lui-même. Il nous apparaît, en effet, que l'apprentissage a pour effet de fournir

¹² Sur ce point, notre étude, rejoint les analyses menées sur les morphogrammes grammaticaux par Béatrix-Kholer (1991).

aux élèves un ensemble de séquences écrites qui sont à leur disposition comme le seraient des éléments d'une seule et même liste.

Hormis un certain nombre de locutions 'prêtes à l'emploi', (par exemple: "c'est", "il y a"), cette liste semble composée d'un ensemble de *mots écrits* au sens strictement graphique du terme, à savoir ce qui, dans la phrase, est écrit entre deux blancs ("tes", "mais", "pour", "déplace", "regardait"). Certains mots semblent avoir fait l'objet d'une appropriation orthographique attestée, mais leur instabilité graphique reste la règle.

Pour ces élèves, l'écriture de la langue consiste alors à exploiter ce stock, en s'attachant transcrire la linéarité phonétique au moyen d'une juxtaposition de mots écrits, sans prise en compte systématique des contraintes qu'impose la morphologie dans le découpage de la chaîne sonore ni de ses effets sur l'orthographe.

L'exemple (16) "ensuite j'ai jouer avec mais jouer" est particulièrement illustratif du procédé utilisé. En effet, c'est par la juxtaposition des mots sélectionnés "avec", "mais", "jouer" que cet élève assure la transcription de l'unité nominale [avɛkmɛʒue]¹³, avec pour point de vue exclusif une correspondance phonie-graphie la plus fidèle possible.

L'écriture des verbes procède de la même manière et les extraits suivants: (18) "se te très très Bien", (19) "c'est ai trô beau, je tes bien la bas"; (20) "ont a regardait" montrent la même stratégie de juxtaposition de mots écrits très étroitement assujettie à la chaîne sonore. On remarquera, en outre, que chacun des mots choisis, lorsqu'il est envisagé pour lui-même, participe bien à écrire des fragments de verbe: bases ou auxiliaires fléchis (ont, a, ai, regardait), pronoms personnels ou réfléchis (je, se, te). En revanche, rapportées aux verbes qu'elles ont pour projet de transcrire, ces successions de fragments s'affranchissent des règles qui contraignent leurs modalités de co-présence dans le verbe: *ont+a+regardait vs on+a+regardé.

3.2 Une morphologie aléatoire

La succession des mots montre une totale indifférence à la problématique de l'homophonie et des incidences allographiques qui en découlent. Plus exactement, le stock de mots graphiques n'est pas filtré par les contraintes liées à l'organisation morphologique interne des unités nominales ou verbales. La notion de morphologie nous est ici utile pour désigner le mode de construction des noms et des verbes, dans la mesure où il répond à un modèle formel marqué par un principe de solidarité interne. Nous entendons par là que les noms et les verbes peuvent être considérés

¹³

A condition de considérer que l'opposition ouvert/fermé est neutralisée.

comme des unités¹⁴ dont l'organisation interne est contrainte par l'ordre, le nombre et la catégorie des morphèmes et lexème en co-présence¹⁵.

Faire le bon choix parmi tous les allographes disponibles (ex: mes, m'est, m'ai) suppose la mise en œuvre d'une analyse qui distribue chacun au sein de programmes unitaires différents, imposant ainsi les catégories des éléments co-présents. C'est ainsi que "mes" ne peut être suivi d'un lexème assorti des suffixes de flexion verbale, mais qu'il peut commuter avec tous les déterminants du nom; en revanche la présence de "m'ai" impose corrélativement et exclusivement celle de "je"; et la présence de "m'est" et de "il/elle" est réciproquement contrainte.

Les erreurs observées montrent que la stratégie de la juxtaposition développée par les élèves ignore cette dimension de l'analyse de la langue.

En effet, c'est bien parce que l'unité nominale a valeur formelle de programme¹⁶ et requiert un déterminant *entre* la préposition et la base que le choix de l'allographe "mais" ("avec mais jouer") est incorrect et que s'impose l'allographe "mes". Pour les mêmes raisons, ne prenant pas appui sur la programmation verbale la séquence (20) "ont + a + regardait" revient à faire se succéder trois graphies de bases verbales fléchies (deux formes du verbe "avoir", une forme du verbe "regarder").

Il est clair que, pour ces élèves, le problème ne réside pas dans une méconnaissance du verbe ou du nom qu'ils utilisent correctement, mais très précisément dans la façon d'écrire ces noms et ces verbes. Comme chacune des formes écrites apprises n'est pas soumise une analyse de dimension segmentale – définissant ainsi sa place et sa catégorie au sein d'un ensemble solidaire – toutes les formes allographes se valent et sont donc substituables l'une à l'autre de façon totalement aléatoire: "ont/on"; "mais/mes"; "déplace/des places". Or, s'il s'agit bien d'une question liée à l'homophonie – une des difficultés cruciales du français – l'ambiguïté est résolue si on se place du point de vue de la segmentation en unités morphologiques. Dans cet ordre d'analyse, selon l'unité – nominale ou verbale – à laquelle elles appartiennent, les séquences [ɔ̃] [mɛ] [deplas]

¹⁴ A la suite de Bonnet et Barreau (1974: 63-67), nous exploitons la notion d'unité de manière restrictive puisque nous considérons que les liens de solidarité interne au nom et au verbe relèvent de la morphologie et non pas de la syntaxe.

¹⁵ "Unité nominale": pour+les+portes vs "unité verbale": avant+ que + tu + me + les + portes.

¹⁶ Les unités sont à concevoir comme des programmes formels d'engendrement des noms et des verbes. Le programme du nom solidarise les fragments suivants: [Préposition + Déterminant nominal + Base]: [avɛk+mɛ+zue] = avec+mes+jouets. Le programme du verbe impose un nombre plus conséquent de fragments: [Conjonction + déterminant verbal + négation₁ + morphème de complément indirect + morphème de complément direct + auxiliaire fléchi + négation₂ + base verbale au participe passé]: [paRsəkə+ty+nə+mə+lɛz+avɛ+pa+dəmɑde] = parce que + tu + ne + me + les + avais + pas + demandés. On comprend dès lors que la maîtrise du verbe, et de son écriture, soit l'objet de difficultés plus conséquentes.

n'occupent pas la même place, ne correspondent pas au même nombre de fragments et n'appartiennent donc pas à la même catégorie.

Formellement, c'est ainsi que [ɔ̃] entre soit dans la liste de substitution des déterminants verbaux, soit dans la liste des allomorphes de la base du verbe "avoir"¹⁷. De la même façon, soit [deplas] se fragmente alors sur le mode d'une base verbale précédée d'un préfixe de dérivation et se définit comme forme fléchie du verbe déplacer "(dé)place", soit la séquence se fragmente en déterminant + base qui en fait un nom "des places". De ce fait, ces formes homophones ne sont jamais commutables l'une pour l'autre au sein d'une même unité¹⁸.

Un des rôles assignés à l'orthographe est, à la fois, de pointer et de lever cette homophonie segmentale formelle par le jeu des techniques d'écritures: les allographes venant marquer la catégorie, et les blancs graphiques matérialiser les frontières (entre les unités et entre les fragments d'une même unité). En conséquence, c'est faute de s'appuyer de manière adéquate sur une segmentation en unités au niveau de la phrase et sur une analyse correcte des fragments qui les composent que les problèmes réciproques de l'homophonie et de l'allographie restent entiers pour les élèves observés. Dans tous les cas, l'effet est le même et il conduit à un brouillage de ce qui est attendu en matière de marquage des frontières à l'écrit.

Conclusion

Notre réflexion, qui devra être complétée par une étude longitudinale, démontre que la stratégie d'encodage graphique adoptée par ces élèves allophones mobilise des acquis indiscutables. Mais elle montre aussi l'avantage qu'il y aurait à initier les élèves à prendre en compte la dimension morphologique de la langue ici entendue non seulement comme "chaînes d'accords morphologiques" (Jaffre & Bessonnat, 1993), mais comme un principe de constitution de programmes d'engendrement du nom et du verbe en français. Les notions de solidarité interne des unités et de segmentation, qui en est son corollaire, nous paraissent ouvrir des pistes didactiques intéressantes pour nourrir les questionnements du passage de l'oral à l'écrit dans l'apprentissage de la langue.

¹⁷ /ɔ̃/ ≠ /ʒə/ ≠ /ty/ ≠ /il/ ≠ /nu/ ≠ /vu/ et doit donc alors être graphié "on"

/ɔ̃/ ≠ /ɛ/ ≠ /a/ ≠ /avɔ̃/ ≠ /ave/ et doit donc alors être graphié "ont"

¹⁸ /deplas/ vs /deRãʒ/ /détaʃ/ etc. et doit donc alors être, graphié "dé-" et lié à la base (préfixe dérivatif).

/deplas/ vs /meplas/ vs /teplas/ vs /kɛlkəplas/, graphié "des" et détaché de la base (déterminant nominal).

Bibliographie

- Besse H. & Porquier R. (1984): Grammaires et didactique des langues. Paris (Hatier).
- Bétrix-Köhler D. (1991): Dis-moi comment tu orthographies, je te dirai qui tu es. Centre Vaudois de Recherches Pédagogiques.
- Bonnet, J. & Barreau, J. (1974): L'esprit des mots. Traité de linguistique française. Tome 1. Grammaire. Paris (l'Ecole).
- Bourdieu, P. (1982): Ce que parler veut dire. L'économie des échanges linguistiques. Paris (Fayard).
- Catach, N. (1980): L'orthographe française – Traité théorique et pratique. Paris (Nathan).
- Cogis, D. (2002): Comment le genre graphique vient aux enfants. In: Hass, G. (dir): Apprendre, comprendre l'orthographe autrement de la maternelle au lycée, CRDP Bourgogne, 19-42.
- Cuq, J.-P. & Gruca, I. (2005): Cours de didactique du français langue étrangère et seconde. Grenoble (PUG).
- Fayol, M. & Jaffré, J. P. (2008): Orthographier. Paris (PUF).
- Gadet, F. (1997): Le français ordinaire. Paris (Armand Colin).
- Jaffré, J. P. & Bessonnat, D. (1993): Accord ou pas d'accord? Les chaînes morphologiques. In: Garcia Debanc, C. (dir): Ecriture et langue, Pratiques, 77, 25-42.
- Moreau, M.-L. (éd). (1997): Sociolinguistique. Concepts de base. Hayen (Mardaga).

Adresses des auteurs

Saliha AMOKRANE, samokrane52@yahoo.fr
Université d'Alger 2, rue Djamel Eddine El Afghani, Bouzaréah, Alger,
Algérie

Jean-Marc DEFAYS & Frédéric SAENEN, jmdefays@ulg.ac.be &
fsaenen@ulg.ac.be
Université de Liège, Institut Supérieur des Langues Vivantes, Place du 20-
Août, 7, B-4000 Liège (Belgique)

Claire FONDET & Fabrice JEJCIC, cfondet@vjf.cnrs.fr & jejcic@vjf.cnrs.fr
UMR 8589, CNRS-LAMOP, Paris 1 Panthéon Sorbonne
Histoire des systèmes graphiques du français et de ses variétés
Laboratoire de médiévistique occidentale de Paris
7 rue Guy Môquet, F-94801 Villejuif cedex

Danièle GEFFROY KONŠTACKÝ & Sylva NOVÁKOVÁ, danièle.geffroy@uhk.cz
& snovakov@krf.zcu.cz
Section de français, Faculté de Pédagogie, Université Hradec Králové,
Rokitanského 62, 500 03 Hradec Králové, République tchèque &
Fakulta pedagogická ZCU, KRF, Jungmannova 1, 306 14 Plzeň, République
tchèque

Jeanne GONAC'H & Clara MORTAMET, jeanne.gonach@univ-rouen.fr &
clara.mortamet@univ-rouen.fr
IRED, 7 rue Thomas Becket, Université de Rouen, F-76821 Mont Saint
Aignan & DESCILAC, UFR Lettres et sciences humaines, Université de
Rouen, F-76821 Mont Saint Aignan Cedex

Claude GRUAZ, claud.gruaz@wanadoo.fr
5, rue aux Boulangers, F-27240 Avrilly

Marie-Claude LE BOT & Elisabeth RICHARD, Marie-claude.lebot@univ-
rennes2.fr & Elisabeth.richard@univ-rennes2.fr
EA LIDILE 3874, UFR Langues, Université Rennes 2,
Place du Recteur Henri Le Moal, CS 24307, F-35043 Rennes cedex

Daniel LUZZATI, daniel.luzzati@univ-lemans.fr
LIUM, Université du Maine, Avenue Olivier Messiaen, F-72085 Le Mans
Cedex 9

Martha MAKASSIKIS & Jean-Christophe PELLAT,
martha.makassikis@gmail.com & pellat@unistra.fr
UR LILPA (Linguistique, langues, parole), Université de Strasbourg, UFR
Lettres, 22, rue Descartes, F-67084 Strasbourg Cedex

Christian SURCOUF, christian.surcouf@unil.ch
École de français langue étrangère (EFLE), Université de Lausanne,
Faculté des lettres UNIL, Dorigny Anthropole, CH-1015 Lausanne

Comité de lecture pour ce numéro

Emmanuelle Carette (CRAPEL, Université de Nancy 2), Francis Carton (CRAPEL, Université de Nancy 2), Virginie Conti (Université de Neuchâtel), Noël Cordonier (Haute Ecole pédagogique, Lausanne), Gilles Corminboeuf (Université de Neuchâtel), Charlotte Dejean (Université de Grenoble 3), Daniel Elmiger (Université de Genève), Anne Grobet (Université de Genève), Jean-Marc Luscher (Université de Genève), Marinette Matthey (Université de Grenoble 3), Jean-François de Pietro (Institut de Recherche et de Documentation Pédagogique, Neuchâtel), Isabelle Racine (Université de Genève), Bertrand Sthioul (Université de Genève), Martine Wirthner (Institut de Recherche et de Documentation Pédagogique, Neuchâtel), Françoise Zay (Université de Genève)